

SORGHUM AND PEARL MILLET AS CLIMATE RESILIENT CROPS FOR FOOD AND NUTRITION SECURITY

EDITED BY: Mahalingam Govindaraj, Wolfram Weckwerth,
Palak Chaturvedi and Velu Govindan

PUBLISHED IN: Frontiers in Plant Science





frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-83250-143-6

DOI 10.3389/978-2-83250-143-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

SORGHUM AND PEARL MILLET AS CLIMATE RESILIENT CROPS FOR FOOD AND NUTRITION SECURITY

Topic Editors:

Mahalingam Govindaraj, HarvestPlus - Alliance Bioversity- CIAT., Colombia

Wolfram Weckwerth, University of Vienna, Austria

Palak Chaturvedi, University of Vienna, Austria

Velu Govindan, International Maize and Wheat Improvement Center (Mexico), Mexico

Citation: Govindaraj, M., Weckwerth, W., Chaturvedi, P., Govindan, V., eds. (2022). Sorghum and Pearl Millet as Climate Resilient Crops for Food and Nutrition Security. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-143-6

Table of Contents

- 06 Editorial: Sorghum and Pearl Millet as Climate Resilient Crops for Food and Nutrition Security**
Palak Chaturvedi, Mahalingam Govindaraj, Velu Govindan and Wolfram Weckwerth
- 10 Wild Sorghum as a Promising Resource for Crop Improvement**
Galaihalage K. S. Ananda, Harry Myrans, Sally L. Norton, Roslyn Gleadow, Agnelo Furtado and Robert J. Henry
- 24 Characterization, Genetic Analyses, and Identification of QTLs Conferring Metabolic Resistance to a 4-Hydroxyphenylpyruvate Dioxygenase Inhibitor in Sorghum (*Sorghum bicolor*)**
Balaji Aravindhan Pandian, Aruna Varanasi, Amaranatha R. Vennapusa, Rajendran Sathishraj, Guifang Lin, Mingxia Zhao, Madison Tunnell, Tesfaye Tesso, Sanzhen Liu, P. V. Vara Prasad and Mithila Jugulam
- 38 Genome-Wide DArTSeq Genotyping and Phenotypic Based Assessment of Within and Among Accessions Diversity and Effective Sample Size in the Diverse Sorghum, Pearl Millet, and Pigeonpea Landraces**
Victor Allan, Mani Vetriventhan, Ramachandran Senthil, S. Geetha, Santosh Deshpande, Abhishek Rathore, Vinod Kumar, Prabhat Singh, Surender Reddymalla and Vânia C. R. Azevedo
- 58 Breeding Strategies and Challenges in the Improvement of Blast Disease Resistance in Finger Millet. A Current Review**
Wilton Mbinda and Hosea Masaki
- 70 Physiological and Proteomic Signatures Reveal Mechanisms of Superior Drought Resilience in Pearl Millet Compared to Wheat**
Arindam Ghatak, Palak Chaturvedi, Gert Bachmann, Luis Valledor, Živa Ramšak, Mitra Mohammadi Bazargani, Prasad Bajaj, Sridharan Jegadeesan, Weimin Li, Xiaoliang Sun, Kristina Gruden, Rajeev K. Varshney and Wolfram Weckwerth
- 94 Cabo Verde's Poaceae Flora: A Reservoir of Crop Wild Relatives Diversity for Crop Improvement**
Vanézia Rocha, Maria C. Duarte, Silvia Catarino, Ivani Duarte and Maria M. Romeiras
- 111 Genetic Variation for Nitrogen Use Efficiency Traits in Global Diversity Panel and Parents of Mapping Populations in Pearl Millet**
Vijayalakshmi Pujarula, Madhu Pusuluri, Srikanth Bollam, Roma Rani Das, Rambabu Ratnala, Gopikrishna Adapala, Vishnukiran Thuraga, Abhishek Rathore, Rakesh K. Srivastava and Rajeev Gupta
- 128 Genetic Gains in Pearl Millet in India: Insights Into Historic Breeding Strategies and Future Perspective**
Om Parkash Yadav, S. K. Gupta, Mahalingam Govindaraj, Rajan Sharma, Rajeev K. Varshney, Rakesh K. Srivastava, A. Rathore and Rajendra Singh Mahala
- 143 Understanding the Sorghum–Colletotrichum sublineola Interactions for Enhanced Host Resistance**
Kibrom B. Abreha, Rodomiro Ortiz, Anders S. Carlsson and Mulatu Geleta

- 157 Nitrogen Use Efficiency in Sorghum: Exploring Native Variability for Traits Under Variable N-Regimes**
Srikanth Bollam, Kirandeep Kaur Romana, Laavanya Rayaprolu, Anilkumar Vemula, Roma Rani Das, Abhishek Rathore, Prasad Gandham, Girish Chander, Santosh P. Deshpande and Rajeev Gupta
- 176 Deciphering Genotype-By-Environment Interaction for Target Environmental Delineation and Identification of Stable Resistant Sources Against Foliar Blast Disease of Pearl Millet**
S. Mukesh Sankar, S. P. Singh, G. Prakash, C. Tara Satyavathi, S. L. Soumya, Yashpal Yadav, L. D. Sharma, A. R. Rao, Nirupma Singh and Rakesh K. Srivastava
- 194 Multi-Environment Quantitative Trait Loci Mapping for Grain Iron and Zinc Content Using Bi-parental Recombinant Inbred Line Mapping Population in Pearl Millet**
Tripti Singhal, C. Tara Satyavathi, S. P. Singh, Aruna Kumar, S. Mukesh Sankar, C. Bhardwaj, M. Mallik, Jayant Bhat, N. Anuradha and Nirupma Singh
- 212 The Sorghum Grain Mold Disease Complex: Pathogens, Host Responses, and the Bioactive Metabolites at Play**
Arlyn Ackerman, Anthony Wenndt and Richard Boyles
- 235 Identifying Anti-Oxidant Biosynthesis Genes in Pearl Millet [*Pennisetum glaucum* (L.) R. Br.] Using Genome—Wide Association Analysis**
Chandra Bhan Yadav, Jayanti Tokas, Devvart Yadav, Ana Winters, Ram B. Singh, Rama Yadav, Prakash I. Gangashetty, Rakesh K. Srivastava and Rattan S. Yadav
- 248 Performance and Stability of Pearl Millet Varieties for Grain Yield and Micronutrients in Arid and Semi-Arid Regions of India**
P. Sanjana Reddy, C. Tara Satyavathi, Vikas Khandelwal, H. T. Patil, P. C. Gupta, L. D. Sharma, K. D. Mungra, Sumer P. Singh, R. Narasimhulu, H. H. Bhadarge, K. Iyanar, M. K. Tripathi, Devvart Yadav, Ruchika Bhardwaj, A. M. Talwar, V. K. Tiwari, U. G. Kachole, K. Sravanti, M. Shanthi Priya, B. K. Athoni, N. Anuradha, Mahalingam Govindaraj, T. Nepolean and Vilas A. Tonapi
- 264 Sorghum Pan-Genome Explores the Functional Utility for Genomic-Assisted Breeding to Accelerate the Genetic Gain**
Pradeep Ruperao, Nepolean Thirunavukkarasu, Prasad Gandham, Sivasubramani Selvanayagam, Mahalingam Govindaraj, Baloua Nebie, Eric Manyasa, Rajeev Gupta, Roma Rani Das, Damaris A. Odeny, Harish Gandhi, David Edwards, Santosh P. Deshpande and Abhishek Rathore
- 281 Genome-Wide Analysis of MADS-Box Genes in Foxtail Millet (*Setaria italica* L.) and Functional Assessment of the Role of SiMADS51 in the Drought Stress Response**
Wan Zhao, Li-Li Zhang, Zhao-Shi Xu, Liang Fu, Hong-Xi Pang, You-Zhi Ma and Dong-Hong Min
- 302 Variation for Photoperiod and Temperature Sensitivity in the Global Mini Core Collection of Sorghum**
Hari D. Upadhyaya, M. Vetriventhan and Vania C. R. Azevedo

- 317 *Harnessing Sorghum Landraces to Breed High-Yielding, Grain Mold-Tolerant Cultivars With High Protein for Drought-Prone Environments***
Mallela Venkata Nagesh Kumar, Vittal Ramya, Mahalingam Govindaraj, Chanda Venkata Sameer Kumar, Setaboyine Maheshwaramma, Seshu Gokenpally, Mathyam Prabhakar, Hariprasanna Krishna, Mulinti Sridhar, Maparla Venkata Ramana, Kodari Avil Kumar and Rumandla Jagadeeshwar
- 338 *Genotypic Variation in Cultivated and Wild Sorghum Genotypes in Response to Striga hermonthica Infestation***
Nicoleta Muchira, Kahiu Ngugi, Lydia N. Wamalwa, Millicent Avosa, Wiliter Chepkorir, Eric Manyasa, Desterio Nyamongo and Damaris A. Odeny
- 354 *Differential Modulation of Heat-Inducible Genes Across Diverse Genotypes and Molecular Cloning of a sHSP From Pearl Millet [Pennisetum glaucum (L.) R. Br.]***
S. Mukesh Sankar, C. Tara Satyavathi, Sharmistha Barthakur, Sumer Pal Singh, C. Bharadwaj and S. L. Soumya
- 375 *Development of Sorghum Genotypes for Improved Yield and Resistance to Grain Mold Using Population Breeding Approach***
C. Aruna, I. K. Das, P. Sanjana Reddy, R. B. Ghorade, A. R. Gulhane, V. V. Kalpande, S. T. Kajjdoni, N. G. Hanamaratti, S. N. Chattannavar, Shivaji Mehtre, Vikram Gholve, K. R. Kamble, C. Deepika, N. Kannababu, D. M. Bahadure, Mahalingam Govindaraj and V. A. Tonapi
- 394 *Genome-Wide Assessment of Population Structure and Genetic Diversity of the Global Finger Millet Germplasm Panel Conserved at the ICRISAT Genebank***
C. Backiyalakshmi, Mani Vetriventhan, Santosh Deshpande, C. Babu, V. Allan, D. Naresh, Rajeev Gupta and Vania C. R. Azevedo
- 408 *Applications of Multi-Omics Technologies for Crop Improvement***
Yaodong Yang, Mumtaz Ali Saand, Liyun Huang, Walid Badawy Abdelaal, Jun Zhang, Yi Wu, Jing Li, Muzafar Hussain Sirohi and Fuyou Wang
- 430 *Pearl Millet: A Climate-Resilient Nutricereal for Mitigating Hidden Hunger and Provide Nutritional Security***
C. Tara Satyavathi, Supriya Ambawat, Vikas Khandelwal and Rakesh K. Srivastava
- 448 *Identification of High-Yielding Iron-Biofortified Open-Pollinated Varieties of Pearl Millet in West Africa***
Prakash I. Gangashetty, Mohammed Riyazaddin, Moussa Daouda Sanogo, Drabo Inousa, Kassari Ango Issoufou, Peter A. Asungre, Ousmane Sy, Mahalingam Govindaraj and Angarawai Ijantiku Ignatius
- 458 *Understanding Heterosis, Genetic Effects, and Genome Wide Associations for Forage Quantity and Quality Traits in Multi-Cut Pearl Millet***
Ponnaiah Govintharaj, Marappa Maheswaran, Michael Blümmel, Pichaikannu Sumathi, Anil Kumar Vemula, Abhishek Rathore, Selvanayagam Sivasubramani, Sandip Mallikarjun Kale, Rajeev Kumar Varshney and Shashi Kumar Gupta



Editorial: Sorghum and Pearl Millet as Climate Resilient Crops for Food and Nutrition Security

Palak Chaturvedi^{1*}, Mahalingam Govindaraj², Velu Govindan³ and Wolfram Weckwerth^{1,4*}

¹ Molecular Systems Biology Lab (MOSYS), Department of Functional and Evolutionary Ecology, Faculty of Life Sciences, University of Vienna, Vienna, Austria, ² Department HarvestPlus Program, Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), Cali, Colombia, ³ International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico, ⁴ Vienna Metabolomics Center (VIME), University of Vienna, Vienna, Austria

Keywords: genetic resources, sorghum, pearl millet, climate smart crops, food and nutritional security, finger millet, foxtail millet, breeding

Editorial on the Research Topic

Sorghum and Pearl Millet as Climate Resilient Crops for Food and Nutrition Security

Sorghum and pearl millet serve as a major source of food, feed and fodder for the semi-arid tropical regions of developing world. These two cereal crops rank within the major six cereal crops with a staple food for about 250 million people residing in semi-arid tropic and dryland areas of south Asia and Africa. Sorghum and pearl millet are also regarded as climate-smart crops because of their extreme tolerance to heat (up to 42°C air temperature), drought, and salinity. This Research Topic on Sorghum and Pearl millet but also Finger millet and Foxtail millet as climate resilient and nutrition-rich crops for food and nutrition security comprise 27 manuscripts. It aims to provide new insights into the genetic resources, high throughput precision phenotyping, breeding approaches, multiomics platforms, gene editing, disease resistance, and gene mapping. It also aims to accelerate breeding cycles for climate resilience and improve nutritional quality in these staple cereal crops.

SORGHUM

In the scope of food and nutritional security, understanding wild progenitors of sorghum (*Sorghum bicolor*) would allow us to exploit the underutilized gene pool to develop more climate-resilient sorghum cultivars. The gene pool of natural sorghum ecotypes may harbor useful gene candidates for both biotic and abiotic stress. Genetic barriers in gene introgression from wild relatives to cultivated sorghum species hold a great challenge. Still, with the recent advent of next-generation sequencing (NGS), more genomic data are available, which expands and extend the sorghum improvement programs using the novel, yet unexploited genes in sorghum's wild relatives (Ananda et al.). Temperature sensitivity and photoperiod of sorghum germplasm are important factors to identify accurate sources for developing cultivars with a broad adaptation, the photoperiod and temperature insensitive, photoperiod and temperature-sensitive and photoperiod sensitive and temperature insensitive sources identified in one of the studies could help breeders to use exact sources in their breeding program, the photoperiod and temperature insensitive accessions can be utilized to develop cultivars with broader adaptation. In contrast, the highly photoperiod sensitive tall accessions can be utilized for biomass and forage improvement and such breeding is suitable in India and USA (Upadhyaya et al.). This segment of research needs more of specific product profile including stress tolerance to meet the special market demand. A major challenge

OPEN ACCESS

Edited and reviewed by:

Sean Mayes,
University of Nottingham,
United Kingdom

*Correspondence:

Palak Chaturvedi
palak.chaturvedi@univie.ac.at
Wolfram Weckwerth
wolfram.weckwerth@univie.ac.at

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 10 January 2022

Accepted: 10 February 2022

Published: 11 March 2022

Citation:

Chaturvedi P, Govindaraj M,
Govindan V and Weckwerth W (2022)
Editorial: Sorghum and Pearl Millet as
Climate Resilient Crops for Food and
Nutrition Security.
Front. Plant Sci. 13:851970.
doi: 10.3389/fpls.2022.851970

in sorghum breeding is the post-emergence grass weed. 4-hydroxyphenylpyruvate dioxygenase-inhibitor herbicides (e.g., mesotrione or tembotrione) can control a broad spectrum of weeds. The sequencing of 317 sorghum lines and QTL mapping genotypes G-200 and G-350 conferred a very high level of metabolic resistance to tembotrione controlled by a polygenic trait (Pandian et al.). Anthracnose is another devastating fungal biotic stress in sorghum caused by *Colletotrichum saublineola*; a review presented by Abreha and coworkers provides a comprehensive overview of the current knowledge on the mechanisms of sorghum-*C. saublineola* molecular interactions, quantitative trait loci (QTL), and major (R) resistance gene sequences as well as defense-related genes associated with anthracnose resistance (Abreha et al.). A systematic validation of these identified genes and QTLs in coming years can assist in breeding resilient sorghum cultivars for stress prone regions particularly in India and Africa.

The contemporary and updated perspective in understanding the genetic and biochemical interactions between the fungal pathogens, their corresponding mycotoxins, and their host has been reviewed (Ackerman et al.). In a multi-location field study, grain yield (GY) and grain mold resistance was tested. Genotype-by-trait biplots indicated that GY is highly influenced by flowering time, 100-grain weight (HGW), and plant height (PH). In contrast, grain mold resistance was influenced by glume coverage and pH (Aruna et al.). Another important parasitic weed in sub-Saharan Africa is *Striga hermonthica*, it is one of the most devastating factors for sorghum production. To identify new sources of resistance to *Striga*, in total 64 sorghum genotypes consisting of landraces, wild relatives, improved varieties, and fourth filial generation (F4) progenies were evaluated for both pot and field trial which resulted in more resistant and high-yielding genotypes from F4 derivatives. These genotypes need more acceptance by the farmers (Muchira et al.). Developing drought-tolerant sorghum varieties with high protein content and tolerance to grain mold is highly important. Nagesh et al. identified four sorghum varieties PYPS 2, PYPS 4, PYPS 8, and PYPS 11, which are highly stable in low grain mold incidence (Kumar et al.). This study used additive main effects, multiplicative interaction (AMMI) and genotype \times environment interaction (GGE) biplot methods.

Pan-genome analysis of sorghum using reference genomes and 354 genetically diverse sorghum accessions led to the identification of more than two million SNPs; association analysis identified approximately 398 SNPs significantly associated with important agronomic traits. Gene expression analysis under drought identified 1,788 genes that were functionally linked to the cell membrane, catalytic activity, molecular function regulation, response to the stimulus, metabolic process, cellular, and biological regulation. In total, 79 genes were absent from the reference genome assembly (Ruperao et al.). More such research analyses are required to strengthen sorghum pan-genome assembly for increased traits association and its use in breeding program.

Improved Nitrogen Use Efficiency (NUE) is one of the primary goals for the global sorghum improvement programs. Root tissues of contrasting lines exhibited differential

expression profiles for transporter genes such as ammonium transporter (SbAMT), nitrate transporters (SbNRT); primary assimilators [glutamine synthetase (SbGS)], glutamate synthase (SbGOGAT[NADH], SbGOGAT[Fd]), assimilatory genes nitrite reductase (SbNiR[NADH]3); and amino acid biosynthesis associated gene [glutamate dehydrogenase (SbGDH)]. Expression profiling of contrasting sorghum genotypes in varying N dosages provides new information in understanding the response of NUE genes toward adaptation to the differential N regimes in sorghum (Bollam et al.). Investigating the biological linkage between and among NUE, stay green and late flowering can offer appropriate breeding road maps for developing optimal NUE in stay green sorghum cultivars in future.

PEARL MILLET

Pearl millet (*Pennisetum glaucum*) breeding in India has historically evolved from open-pollinated varieties to single cross hybrid breeding in a comprehensive manner with closer and continued association of CGIAR and NARS centers. To further accelerate the hybrid breeding efforts for drought-prone areas in South Asia and Sub-Saharan Africa, the heterotic grouping of hybrid parental lines is essential to sustain long-term genetic gains (Yadav O. P. et al.). Pearl millet is nutritionally rich and high in micronutrients such as iron (Fe) and zinc (Zn) and its increased dietary intake can prevent associated hidden hunger or malnutrition. The inclusion of minimum standards for micronutrients such as Fe and Zn content in the cultivar grain release policy is for the first time reported in pearl millet across the globe, motivate institutional commitments and progress toward incorporating essential nutritional traits in breeding pipelines (Satyavathi et al.). QTLs for Fe and Zn content from three distinct production environments were generated using a genetic linkage map consisting of 210 F6 recombinant inbred lines (RIL) population derived from the (PPMI 683 \times PPMI 627) cross using genome-wide simple sequence repeats (SSRs). Two constitutive expressing QTLs for Fe and Zn were co-mapped in LG 2. The second one on LG 3, the QTLs candidate genes such as Ferritin gene, Al^{3+} , K^{+} , Zn^{2+} and Mg^{2+} transporters were identified using bioinformatics approaches (Singhal et al.). In another study, newly developed open-pollinated varieties (30 OPVs of which 8 are Fe/Zn biofortified) were tested for field performance and stability for grain yield, grain Fe and Zn contents across 10 locations in West Africa, resulting in a strong correlation ($r = 0.98^{**}$) between grain Fe and Zn contents that merit Fe-based selection and can be effective in pearl millet variety breeding (Gangashetty et al.).

Importance of open pollinated varieties cannot be ruled out because of lower input cost, wider adaptation and timely seed availability. OPVs of pearl millet were tested in three different locations across India to check the variation in grain Fe and Zn contents. The results showed a highly significant positive correlation (across environment = 0.83; $p < 0.01$), indicating the efficacy of simultaneous selection for both traits (Sanjana Reddy et al.). A set of 105 forage-type hybrid parents of the diverse panel was genotyped following genotyping by sequencing

(GBS) and phenotyped for crude protein (CP) under multi-cuts for two consecutive years. This led to the identification of one stable significant single nucleotide polymorphism (SNP) on LG4 for CP. Nine SNPs were distributed across six linkage groups except on LG2 (Govintharaj et al.). These identified loci require validation with robust phenotyping methods in forage gene pool including photo sensitive breeding materials which can facilitate forage quality traits improvement in pearl millet through marker-assisted selection.

Transcript expression profiling for functional classification of a gene belonging to a small heat shock protein (sHSP) family in pearl millet under high-temperature stress led to the identification of two high-temperature-responsive markers Pgcp70 and PgHSPF. Physio-biochemical trait screening of the contrasting genotypes among the eight different pearl millet inbred lines at the seedling stage resulted in the identification of PgHSP20 genes, which can provide further insights into the molecular regulation of pearl millet stress tolerance, thereby bridging them together to fight against the unpredicted nature of abiotic stress (Mukesh Sankar, Satyavathi et al.).

Foliar blast disease of pearl millet is severe, caused by *Magnaporthe grisea*. To unravel the G x E interactions for identification and validation of stable resistant genotypes against foliar blast disease through multi-environment testing, a group of 250 different accessions from 20 different countries were collected and screened under natural epiphytotic conditions, which resulted in 43 resistant genotypes which can be used in future resistance breeding programs for pearl millet (Mukesh Sankar, Singh et al.).

Pearl millet accessions that can use nitrogen efficiently needs to be characterized soon. In this aim in total 380 diverse pearl millet lines consisting of a global diversity panel (345), parents of mapping populations (20), and standard checks (15) were evaluated in an alpha-lattice design with two replications. Eleven nitrogen use efficiency (NUE) related traits across three growing seasons in an N-depleted precision field under three different N levels (0%-N0, 50%-N50, 100%-N100 of recommended N, i.e., 100 kg ha⁻¹) resulted in 25 top N-tolerant and N-sensitive genotypes under low N conditions. Tolerant genotypes with low N may help identify genomic regions responsible for NUE. Its deployment in pearl millet breeding programs through marker-assisted selection (MAS) can be facilitated (Pujarula et al.). Cabo Verde Islands are poorly explored for genetic resources related to plants. Their potential to supplement the genetic pool of cultivated species is an attempt to identify islands crop wild relatives (CWR) from the Poaceae family and provide a checklist of priority CWR taxa, highlighting particular conservation concerns and the areas which should be the focus of the most intensive conservation efforts in these islands (Rocha et al.). Similarly, the total antioxidant content of pearl millet flour and evaluation of 222 genotypes for antioxidant activity from inbred lines resulted in 18 candidate genes related to antioxidant pathway genes (flavanone 7-O-beta-glycosyltransferase, GDGL esterase/lipase, glutathione S-transferase) residing within or near the association signal that can be selected for further functional characterization (Yadav C. B. et al.).

Multomics combined with speed breeding is one of the answers to producing highly nutritious food crops (Weckwerth et al., 2020; Yang et al.). Furthermore, integration of the individual omics technique employing the “phenotype to genotype” and “genotype to phenotype” concept together with the systems biology approach may be beneficial for crop breeding improvement under different environmental conditions (Weckwerth et al., 2020). Recently, two important cereal crops, Pearl millet (C₄) and Wheat (C₃), were compared at the physiological and proteomics level to understand the drought stress response mechanisms. Tissue-specific proteome analysis of leaves, roots and seeds led to the identification of 12,558 proteins in pearl millet and wheat under well-watered and stress conditions. The physiological response was demonstrated using Odum's model. The study provides for the first time “stay-green” proteomics signatures for Pearl millet (Ghatak et al.). Furthermore, comparative proteome signatures for “stay-green” and “senescence” traits in Pearl millet and wheat under drought stress were identified and correlated with the physiological analysis. NAD-ME type photosynthesis was evaluated in both the cereals, and discriminant analysis via sPLS led to the identification of the putative protein markers, and correlation with an important physiological trait such as root length was determined. This study provides an opportunity to identify important molecular processes in C₄ traits essential for drought resistance and incorporate them into C₃ plants via genetic engineering (Ghatak et al.).

FINGER MILLET AND FOXTAIL MILLET

The Research Topic also consists of manuscripts on finger millet (*Eleusine coracana*) and foxtail millet (*Setaria italica*), also members of the Poaceae family. Finger millet is an important cereal crop in southern Asia and eastern Africa. It has a long storage period, grows under arid and semi-arid environmental conditions, and has good nutraceutical properties. Blast disease in finger millet caused by the filamentous ascomycetous fungus (*Magnaporthe oryzae*) is the most devastating disease affecting the growth and yield of this crop in all its growing regions. Breeding strategies and challenges in improving this blast disease resistance in finger millet have been extensively reviewed (Mbinda and Masaki). A total of 314 global finger millet germplasm diversity panel accessions were genotyped, using the DArTseq approach to find the genetic diversity and population structure within these genotypes, the authors obtained 33,884 high-quality single nucleotide polymorphism (SNP) markers on 306 accessions after filtering, considerable genetic diversity, and the mean polymorphic information content was determined (Backiyalakshmi et al.). In crops, MADS-box transcription factors play vital roles in multiple biological processes. Genome-wide identification and classification of MADS-box genes in foxtail millet have not been reported previously. In total, 72 MADS-box genes in the foxtail millet genome give an overview of the phylogeny, chromosomal location, gene structures, and potential functions of the proteins encoded by these genes. Expression patterns of 10 foxtail millet MADS-box genes that

are upregulated in response to drought were analyzed in different tissues in response to different abiotic stresses because the SiMADS51 genes were found to be strongly induced by drought stress, the function of the SiMADS51 gene was assessed by expression in the model plants *Arabidopsis* (*Arabidopsis Thaliana*) and rice (*Oryza sativa* L.) (Zhao et al.). In another study, 108 diverse landraces and wild accessions of sorghum, pearl millet, and pigeon pea were studied by genotyping using the DARTSeq approach, which identified 45249 SNPs in pearl millet, 19052 in SNPs sorghum and 8211 SNPs in pigeonpea. Interestingly, sorghum had the lowest average phenotypic (0.090) and genotypic (0.135) variance within accession distances, while pearl millet had the highest average phenotypic (0.227) and genotypic (0.245) distances. These studies are very helpful to the genebank curators to understand the dynamics of the population within accession and support the planning of appropriate germplasm conservation strategies (Allan et al.).

In summary, the variety of studies reported in these diverse crops, pearl millet, sorghum, finger millet and foxtail millet are comprehensive and provides immense knowledge to the coming generations of crop scientists, crop physiologists, plant

biologists and breeders. The studies reported in this Research Topic (Volume I) provide us with clear global research goals and are in place on making more climate-resilient crops in future by close observing crop agro-climate variability. Breeders are provided with specific and comprehensive catalogs of important and validated gene candidates that are associated with resilience and nutritional traits. These groundbreaking studies and corresponding breeding programs will eventually enhance crop productivity and improve sorghum and millet-based food intake to meet the food and nutritional security in south Asia and sub-Saharan Africa. They also open up the path to new exploitation of these prestigious cereal crop plants in other regions of the world subject to climate crisis.

AUTHOR CONTRIBUTIONS

PC and WW drafted the editorial. MG contributed and provided inputs at the designing stage of the Research Topic. PC, MG, VG, and WW reviewed and revised the editorial. All authors approved the submitted version.

REFERENCES

Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P., and Varshney, R. K. (2020). PANOMICS meets germplasm. *Plant. Biotechnol. J.* 18, 1507–1525. doi: 10.1111/pbi.13372

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chaturvedi, Govindaraj, Govindan and Weckwerth. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Wild Sorghum as a Promising Resource for Crop Improvement

Galaihalage K. S. Ananda^{1†}, Harry Myrans^{2†}, Sally L. Norton³, Roslyn Gleadow^{1,2‡}, Agnelo Furtado¹ and Robert J. Henry^{1*}

¹ Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, St Lucia, QLD, Australia, ² School of Biological Sciences, Monash University, Clayton, VIC, Australia, ³ Australian Grains Genebank, Agriculture Victoria, Horsham, VIC, Australia

OPEN ACCESS

Edited by:

Petr Smýkal,
Palacký University, Olomouc, Czechia

Reviewed by:

Maryke T. Labuschagne,
University of the Free State,
South Africa
Shivali Sharma,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

*Correspondence:

Robert J. Henry
robert.henry@uq.edu.au

[†]These authors share first authorship

*ORCID:

Roslyn Gleadow
orcid.org/0000-0003-4756-0411

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 04 May 2020

Accepted: 06 July 2020

Published: 17 July 2020

Citation:

Ananda GKS, Myrans H, Norton SL,
Gleadow R, Furtado A and Henry RJ
(2020) Wild Sorghum as a Promising
Resource for Crop Improvement.
Front. Plant Sci. 11:1108.
doi: 10.3389/fpls.2020.01108

Sorghum bicolor (L.) Moench is a multipurpose food crop which is ranked among the top five cereal crops in the world, and is used as a source of food, fodder, feed, and fuel. The genus *Sorghum* consists of 24 diverse species. Cultivated sorghum was derived from the wild progenitor *S. bicolor* subsp. *verticilliflorum*, which is commonly distributed in Africa. Archeological evidence has identified regions in Sudan, Ethiopia, and West Africa as centers of origin of sorghum, with evidence for more than one domestication event. The taxonomy of the genus is not fully resolved, with alternative classifications that should be resolved by further molecular analysis. Sorghum can withstand severe droughts which makes it suitable to grow in regions where other major crops cannot be grown. Wild relatives of many crops have played significant roles as genetic resources for crop improvement. Although there have been many studies of domesticated sorghum, few studies have reported on its wild relatives. In *Sorghum*, some species are widely distributed while others are very restricted. Of the 17 native sorghum species found in Australia, none have been cultivated. Isolation of these wild species from domesticated crops makes them a highly valuable system for studying the evolution of adaptive traits such as biotic and abiotic stress tolerance. The diversity of the genus *Sorghum* has probably arisen as a result of the extensive variability of the habitats over which they are distributed. The wild gene pool of sorghum may, therefore, harbor many useful genes for abiotic and biotic stress tolerance. While there are many examples of successful examples of introgression of novel alleles from the wild relatives of other species from Poaceae, such as rice, wheat, maize, and sugarcane, studies of introgression from wild sorghum are limited. An improved understanding of wild sorghums will better allow us to exploit this previously underutilized gene pool for the production of more resilient crops.

Keywords: sorghum, crop wild relatives, crop improvement, cyanogenesis, wild sorghum

INTRODUCTION

Sorghum bicolor (L.) Moench, commonly known as sorghum, is ranked among the five main cereal crops in the world (Mace et al., 2009; Venkateswaran et al., 2014). It plays a vital role in global food production and is the staple food of billions of people (Mace et al., 2009). Sorghum is a multipurpose crop cultivated for grain, sweet stem, forage, and broomcorn. It also serves as a source of fuel,

bioethanol, alcoholic beverages, and building materials. It is one of the most important food crops of arid and semi-arid regions of the world, whereas in developed countries it is grown mainly for forage and animal feed (Hariprasanna and Patil, 2015; Venkateswaran et al., 2019a). Currently, the USA has the world's greatest total sorghum production, followed by Nigeria, India, and Mexico, with an average global production of 50 megatons per year (FAOSTAT, 2019). Sorghum is well adapted to high temperature, dry conditions and it is surprising that it is not even more widely grown. The lower global production of sorghum, relative to other cereals such as wheat and rice, might be increased by the exploitation of the hitherto untapped potential of the extensive gene pool of crop wild relatives (CWR) in the genus (Sasaki and Antonio, 2009). The use of *Sorghum* genetic resources is most immediately applicable to production of improved sorghum varieties. *Sorghum* is a genus within the tribe, Andropogoneae that includes other genera of plants such as *Saccharum* (Bonnett and Henry, 2011) and *Miscanthus* (Anzoua et al., 2011) that are important biomass crops. Sugarcane (*Saccharum*) and sorghum (*Sorghum*) are closely related and may be inter-crossed (Gupta et al., 1978). CWR in the *Sorghum* genus may also be a genetic resource to support the development of new crops across the tribe either by introgression of useful genes into genera such as *Saccharum* or by domestication of further *Sorghum* species (Dillon et al., 2007c).

CWR are plant species that are closely related to a domesticated crop, from anywhere in the world, including crop progenitors, landraces, and closely related taxa not historically involved in agriculture. They represent one of the key sources of new genetic material to introduce to crop lines through traditional breeding and, to a lesser extent, genetically modified (GM) crops. The use of CWR by agricultural scientists started to become a regular practice in the 1940s (Meilleur and Hodgkin, 2004). They have since been used to produce new lines of many globally important crops, improving traits such as disease and pest resistance, nutritional value, yield, and tolerance to abiotic stresses in crops such as wheat, tomatoes, rice, and many others (Prescott-Allen and Prescott-Allen, 1986; Hajjar and Hodgkin, 2007). CWR are seen by many as an invaluable source of diversity which should be drawn upon to further enhance crops in terms of commercial value, and to facilitate adaptation to changing environments and pathogens (Hoyt, 1988; Jarvis et al., 2008; Dempewolf et al., 2014; Brozynska et al., 2016). In monetary terms, it is estimated that the genetic resources they are worth over US\$150bn (Tyack et al., 2015), highlighting the vital role they could potentially play in agriculture. Here, we refer to sorghum's CWR as the wild taxa in the genus *Sorghum* Moench, including sorghum's progenitors, but not landraces. Species names and ranks were standardized according to the USDA (2020). This review aims to understand the historical and current uses of sorghum crops and difficulties facing sorghum agriculture, and explores CWR's potential as viable resources for future genetic improvement of the crop. To do this we discuss the origins and domestication of the crop, summarize and clarify what is known of the taxonomy of the genus and the

phylogenetic relationships between subgenera, the barriers to gene flow and the potential for crop improvement.

ORIGINS AND DISTRIBUTION OF DOMESTICATED SORGHUM

The earliest evidence of use of wild sorghum as a food is from the Sahara, around 7500 BC, where hunter-gatherers lived (Venkateswaran et al., 2019a). Similarly, a recent study by Winchell et al. (2017), has shown that the earliest domesticated sorghums are found in Neolithic populations of Sudan around fourth millennium BC. The exact origin and location of sorghum domestication is debated (De Wet et al., 1970; Venkateswaran et al., 2019a), however, archaeological evidence supports domestication in eastern Sudan around 3000 BC (Fuller and Stevens, 2018). Some studies suggest that there may have been more than one domestication event, potentially explaining the origin of the group *guinea-margaritifera* of genus *Sorghum*, which was domesticated more recently (Kimber, 2000; Mace et al., 2013). According to archaeological evidence, *S. bicolor* originated from its wild progenitor *Sorghum bicolor* (L.) Moench subsp. *verticilliflorum* (Steud.) de Wet ex Wiersema & J. Dahlb., which is commonly distributed in Africa (De Wet and Harlan, 1971; De Wet, 1978; Doggett, 1988). There is no direct evidence available to suggest any contribution of other wild relatives viz., *Sorghum propinquum* (Kunth) Hitchc. and *Sorghum halepense* (L.) Pers. to cultivated sorghum, as suggested by Doggett (1988). Rowley-Conwy et al. (1997), proposed three hypotheses for sorghum domestication. The first hypothesis is based on the studies of Murdock (1959), which described an independent nuclear Mande center in West Africa. The next hypothesis is that the origin of sorghum could be in eastern Sahara, around 9700–6200 BC (Ehret, 2014), and the final hypothesis relies on the evidence of the race *durra* in India back in 4000 BC.

From its first ancestor in Africa, domesticated sorghum was distributed across the globe by various means—most commonly along trade routes. From East Africa, cultivated sorghum was moved across eastern and southern Africa as a result of human migration (Mann et al., 1983). It was then introduced to India via the Middle East trade routes (Mann et al., 1983). Doggett (1988), reported overland routes from East Africa and Somalia via Aden. The earliest *Sorghum* species found in India was *S. bicolor* and evidence for domestication and cultivation dating back to c.2000–1700 BC was found in the Indus Valley (Meadow, 1996; Fuller, 2003). Since then, sorghum has played a key role in agriculture in India (Kleih et al., 2000) and India is now considered to be its secondary center of diversity (Appa et al., 1996).

Sorghum was introduced to China from India, again via sea and overland trade routes. There are several hypotheses on how sorghum arrived in China. One of the possible ways was through the river valleys of Indochina (Venkateswaran et al., 2019a). However, Hagerty (1941) claims that the emperor Genghis Khan introduced sorghum to China after his voyage to South Asia between AD 1206–1228. The *Amber cane sorghos* are related to

eastern African sorghums whereas the *Kaoliangs* probably originated from the *Sorghum bicolor* introduced from India (Doggett, 1988). There is evidence that *Kaoliangs* might be derived from native wild diploid sorghum (Harlan, 1995). The Yellow River Valley is considered to be the area where the earliest sorghum was cultivated based on archaeological evidence (Venkateswaran et al., 2019a). From China, sorghum was brought to the USA by the slave traders in the 19th century. According to Martin (1936), the first sorghum to be introduced to the USA was the Chinese *Amber* in 1853. Sorghum was introduced to Queensland, Australia, in the 1900s by Americans (Venkateswaran et al., 2014). Since then sorghum has become a major summer crop in Australia, accounting for 5% of the global export of sorghum globally (Venkateswaran et al., 2019a).

TAXONOMY OF SORGHUM

The genus *Sorghum* was first classified as *Holcus* by Linnaeus in 1753 and constituted three species; *Holcus sorghum*, *Holcus saccharatus* and *Holcus tricolor*. *Sorghum* was separated out from the genus *Holcus* by Moench in 1794 (Venkateswaran et al., 2019a). Following these classifications, domesticated sorghum was formally recognized as *Sorghum bicolor* (L.) Moench (Venkateswaran et al., 2014; Hariprasanna and Patil, 2015). According to the current classification, sorghum belongs to the kingdom Plantae, division Magnoliophyta, class Liliopsida, order Cyperales, family Poaceae, tribe Andropogoneae, subtribe

Sorghinae, and genus *Sorghum* (Hariprasanna and Patil, 2015). In Snowden's classification, sorghum was divided into two main sections, *Eu-sorghum* and *Parasorghum*, based on morphological traits such as color of grains and glumes and persistence of pedicellate spikelets (Snowden, 1955). However, five subgenera of *Sorghum* are now recognized: *Eu-sorghum*, *Chaetosorghum*, *Heterosorghum*, *Parasorghum*, and *Stiposorghum* (Garber and Snyder, 1951; Harlan and de Wet, 1972; De Wet, 1978; Lazarides et al., 1991), based on morphological characters (**Figure 1**). Despite *S. bicolor* having been domesticated in East Africa, 17 *Sorghum* species are native to Australia. Of these, 13 are endemic, emphasizing the need to preserve sorghum's CWR nationally. Native Australian species are present in every *Sorghum* subgenus, excepting *Eu-sorghum*.

The exact number of species in this highly diverse genus is still not well established. According to Dillon et al. (2001), *Sorghum* consists of 25 species distributed across Australia, the Pacific Islands, Southeast, East and South Asia, and much of Africa (**Table 1**). The USDA recently accepted one additional species to the genus—*Sorghum trichocladium* (Rupr. ex Hack.) Kuntze, which is native to Mexico, Guatemala, and Honduras (USDA, 2020). This species can be found only in limited locations (Spangler, 2003) and limited information is available on this species. Kew's Angiosperm DNA C-values database, however, currently lists a total of 32 *Sorghum* species (Leitch et al., 2019). These differing classifications are based on diverse parameters, making sorghum taxonomy a complex and debatable area of study. In this review, *Sorghum* consists of 24 accepted species (USDA, 2020), with *S. bicolor* subsp.

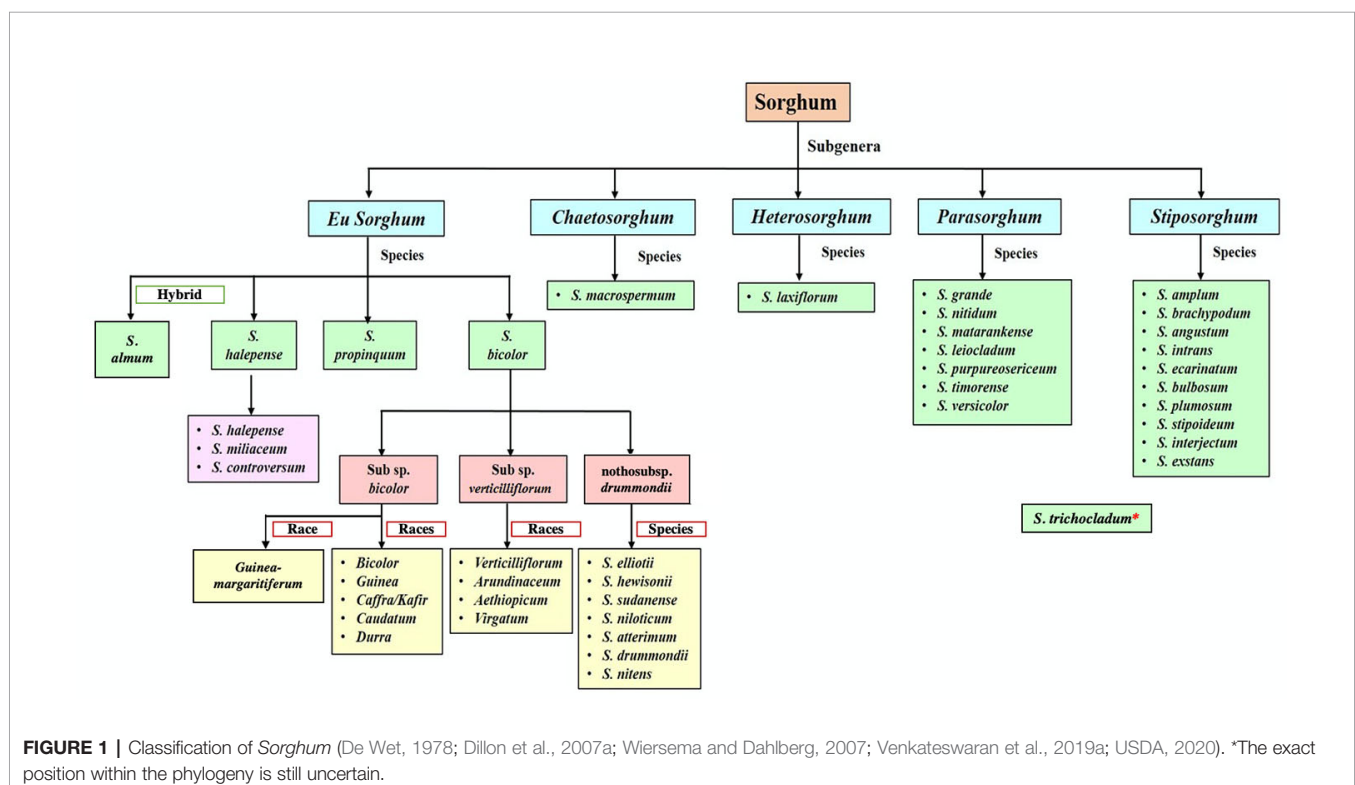


TABLE 1 | Taxonomic information, life form, and ploidy levels of taxa in the genus *Sorghum*.

Taxon and subgeneric section	Subgeneric section	Gene pool	Current Accepted taxonomy (USDA) (in AGG Grin Global database to be live Nov 2019)	Lifeform/Duration	Ploidy (2n)
<i>S. xalmum</i> Parodi	<i>Eu-sorghum</i>	Secondary	<i>S. xalmum</i> Parodi	Perennial	40
<i>S. arundinaceum</i> (Desv.) Stapf	<i>Eu-sorghum</i>	Primary	<i>Sorghum bicolor</i> subsp. <i>verticilliflorum</i> (Steud.) de Wet ex Wiersma & J. Dahlb.	Annual	20
<i>S. bicolor</i> (L.) Moench	<i>Eu-sorghum</i>	Primary	<i>S. bicolor</i> (L.) Moench	Annual	20
<i>S. xdrummondii</i> (Steud.) Millsp. & Chase	<i>Eu-sorghum</i>	Primary	<i>Sorghum bicolor</i> nothosubsp. <i>drummondii</i> (Steud.) de Wet ex Davidse	Annual	20
<i>S. halepense</i> (L.) Pers.	<i>Eu-sorghum</i>	Secondary	<i>S. halepense</i> (L.) Pers.	Perennial	40
<i>S. propinquum</i> (Kunth) Hitchc.	<i>Eu-sorghum</i>	Primary	<i>S. propinquum</i> (Kunth) Hitchc.	Perennial	20
<i>S. grande</i> Lazarides	<i>Parasorghum</i>	Tertiary	<i>S. grande</i> Lazarides	Perennial	30, 40
<i>S. leiocladum</i> (Hack.) C. E. Hubb.	<i>Parasorghum</i>	Tertiary	<i>S. leiocladum</i> (Hack.) C. E. Hubb.	Perennial	10, 20
<i>S. matarankense</i> E. D. Garber & Snyder	<i>Parasorghum</i>	Tertiary	<i>S. matarankense</i> E. D. Garber & Snyder	Perennial	10
<i>S. nitidum</i> (Vahl) Pers.	<i>Parasorghum</i>	Tertiary	<i>S. nitidum</i> (Vahl) Pers.	Perennial	10, 20
<i>S. purpureosericeum</i> (Hochst. ex. A. Rich.) Asch. & Schweinf.	<i>Parasorghum</i>	Tertiary	<i>S. purpureosericeum</i> (Hochst. ex. A. Rich.) Asch. & Schweinf.	Annual	10
<i>S. timorensis</i> (Kunth) Buse	<i>Parasorghum</i>	Tertiary	<i>S. timorensis</i> (Kunth) Buse	Perennial	10, 20
<i>S. versicolor</i> Andersson	<i>Parasorghum</i>	Tertiary	<i>S. versicolor</i> Andersson	Annual	10, 20
<i>S. amplum</i> Lazarides	<i>Stiposorghum</i>	Tertiary	<i>S. amplum</i> Lazarides	Annual	10, 30
<i>S. angustum</i> S. T. Blake	<i>Stiposorghum</i>	Tertiary	<i>S. angustum</i> S. T. Blake	Annual	10
<i>S. brachypodium</i> Lazarides	<i>Stiposorghum</i>	Tertiary	<i>S. brachypodium</i> Lazarides	Annual	10
<i>S. bulbosum</i> Lazarides	<i>Stiposorghum</i>	Tertiary	<i>S. bulbosum</i> Lazarides	Annual	10
<i>S. ecarinatum</i> Lazarides	<i>Stiposorghum</i>	Tertiary	<i>S. ecarinatum</i> Lazarides	Annual	10
<i>S. exstans</i> Lazarides	<i>Stiposorghum</i>	Tertiary	<i>S. exstans</i> Lazarides	Annual	10
<i>S. interjectum</i> Lazarides	<i>Stiposorghum</i>	Tertiary	<i>S. interjectum</i> Lazarides	Annual/ Perennial	30
<i>S. intrans</i> F. Muell. ex Benth.	<i>Stiposorghum</i>	Tertiary	<i>S. intrans</i> F. Muell. ex Benth.	Annual	10
<i>S. plumosum</i> (R. Br.) P. Beauv.	<i>Stiposorghum</i>	Tertiary	<i>S. plumosum</i> (R. Br.) P. Beauv.	Annual	10, 20, 30, 40
<i>S. stipoides</i> (Ewart & Jean White) C. A. Gardner & C. E. Hubb.	<i>Stiposorghum</i>	Tertiary	<i>S. stipoides</i> (Ewart & Jean White) C. A. Gardner & C. E. Hubb.	Annual	10
<i>S. laxiflorum</i> F. M. Bailey	<i>Heterosorghum</i>	Tertiary	<i>S. laxiflorum</i> F. M. Bailey	Annual	40
<i>S. macrospermum</i> E. D. Garber	<i>Chaetosorghum</i>	Tertiary	<i>S. macrospermum</i> E. D. Garber	Annual	40
<i>S. trichocladum</i> (Rupr. ex Hack.) Kuntze	–	Tertiary	<i>S. trichocladum</i> (Rupr. ex Hack.) Kuntze	Perennial	–

verticilliflorum and *drummondii* no longer considered separate species (Figure 1). The traditional classification of the genus, based on morphological parameters (Venkateswaran et al., 2014), is of limited value because it results in significant overlapping of the existing taxa. By contrast, recent studies of sorghum based on molecular evidence, such as phylogenetic analyses of DNA sequencing data, have been able to generate a classification with clear and precise groupings of these species (Venkateswaran et al., 2014). Although weak molecular evidence suggests that *S. trichocladum* is closely related to Australian taxa, the exact position of *S. trichocladum* in the phylogeny remains uncertain (Spangler, 2003).

Eu-sorghum

Eu-sorghum is one of two major sections in the genus *Sorghum*. It is mainly distributed in Africa and southern Asia (Price et al., 2005a). In the original classification by Snowden (1955) *Eu-sorghum* was divided into two sub-sections, *Arundinaceae* and *Halepensis*. The sub-section *Arundinaceae* was further divided into two series, *Spontanea* (grass sorghum) and *Sativa* (grain sorghum). *Spontanea* contained 10 wild species whereas *Sativa* contained 31 cultivated species. Sub-section *Halepensis* was comprised of four wild rhizomatous taxa (De Wet and Harlan, 1971). Subsequently this classification was modified by many scientists. The number of members included in each group

varied with the classification. For instance, the classification of de Wet et al. (1970) placed 17 wild species in the complex of *Spontanea*, while 31 cultivated species were in *Sativa* and the sub-section *Halepensis* contained four wild grass species.

In the currently accepted classification, *Eu-sorghum* is considered the “true sorghum” and contains three species, *S. bicolor*, *S. propinquum*, *S. halepense* and a hybrid species called, *Sorghum xalmum* Parodi (USDA, 2020).

Sorghum bicolor includes most cultivated sorghum lines, and is distinguished from other species by the bulky, open inflorescence and the non-pendulous branches separating at the base. *Sorghum bicolor* can be separated into three subspecies: subsp. *bicolor* (all cultivated sorghums), subsp. *verticilliflorum* (wild progenitors of cultivated sorghums), and nothosubsp. *drummondii* (Steud.) de Wet ex Davidse (weedy hybrids and the derivatives of hybridization between *S. bicolor* subsp. *bicolor* and *verticilliflorum*). The subsp. *verticilliflorum* (formerly known as *arundinaceum*) (Venkateswaran et al., 2019a) consists of four races of wild progenitors: *aethiopicum*, *arundinaceum*, *verticilliflorum*, and *virgatum*. The race *arundinaceum* is distributed mostly in Africa and has a large and exposed inflorescence as well as flexuous branches which are not dividing at the base. The desert grass, race *aethiopicum*, is widely distributed in the African Sahel and has a comparatively small, constricted inflorescence together with divided sub-erect

branches. The race *virgatum*, characterized by a slender inflorescence and narrow, linear leaf blades, is widespread in north eastern Africa. The race *verticilliflorum* which is native to Africa and distributed in Madagascar can be characterized by a large open inflorescences with spreading branches divided at the base (Venkateswaran et al., 2014; Venkateswaran et al., 2019b). In the Snowden (1955) classification, there are seven weedy taxa recognized in nothosubsp. *drummondii* (also known as “Sudan grass”), which are commonly cultivated as forage. The currently accepted five races of subsp. *bicolor* are: *bicolor*, *guinea*, *kafir*, *caudatum*, and *durra*, which are categorized in this subspecies based only on their spikelet morphology with 10 intermediate races (Lazarides et al., 1991). Based on molecular evidence, Mace et al. (2013) also separated *guinea-margaritifera* as a distinct race of subsp. *bicolor*, with this group previously being included in the broader *guinea* race. *Guinea-margaritifera* represents an intermediate race between the wild subsp. *verticilliflorum* and the other domesticated races of subsp. *bicolor*.

Sorghum propinquum is a diploid rhizomatous wild perennial species that is distributed in Southeast Asia and Indian subcontinent. Smaller spikelets are distinctive features of *S. propinquum*. Another perennial species, *S. halepense*, also known as “Johnson grass,” is a tetraploid rhizomatous wild relative that is widespread in Southern Eurasia and India. According to Snowden classification 1955, this species contains members of three former species known as *S. halepense*, *S. miliaceum* (Roxb.) Snowden, and *S. controversum* (Steud.) (Venkateswaran et al., 2014). *Sorghum halepense* has comparatively large inflorescences than other two species. These two wild perennial species have given rise to hybrids and hybrid derivatives as a result of their introgression with *S. bicolor* (Dahlberg, 2000; Venkateswaran et al., 2014). *Sorghum xalmum*, for example, is a hybrid between *S. bicolor* and *S. halepense* (Duvall and Doebley, 1990; Dillon et al., 2007b).

Parasorghum

The section *Parasorghum* includes seven species: *Sorghum grande* Lazarides, *Sorghum leiocladum* (Hack.) C. E. Hubb., *Sorghum matarankense* E. D. Garber & Snyder, *Sorghum nitidum* (Vahl) Pers., *Sorghum purpureosericeum* (Hochst. ex A. Rich.) Schweinf. & Asch., *Sorghum timorense* (Kunth) Büse, and *Sorghum versicolor* Andersson. Excepting *S. purpureosericeum* and *S. versicolor*, all *Parasorghum* species are native to Australia, with *S. grande*, *S. leiocladum* and *S. matarankense* all being endemic (Lazarides et al., 1991). *Sorghum grande* is a perennial diploid with a chromosome number of 30 or 40 ($2n = 30, 40$), distributed in the Northern Territory (isolated in Katherine region) and Queensland, Australia. *Sorghum nitidum* is also a perennial diploid with $2n = 10, 20$ and is found in Queensland, New Guinea, and Southeast and East Asia. *Sorghum leiocladum* is a perennial with $2n = 20$ which is distributed in southern Queensland, New South Wales, and northern Victoria. *Sorghum matarankense* is an annual species with $2n = 10$ and it can be commonly seen in north-central parts of the Northern Territory, Australia. Likewise, *S. timorense* is an annual species with $2n = 10, 20$, found in northern Australia and Timor. *Sorghum timorense* is distinguished by a minute, sessile spikelet with an obovoid caryopsis and a developed pedicellate spikelet. *Sorghum purpureosericeum* is an annual with chromosome number $2n = 10$ and 20, and is found

in India, the Sahel, and east and west tropical Africa. *Sorghum versicolor* is also annual, with a chromosome number of $2n = 10$ and 20 and is found in eastern and southern Africa.

Stiposorghum

The subgenus *Stiposorghum* contains a total of 10 species: *Sorghum amplum* Lazarides, *Sorghum brachypodium* Lazarides, *Sorghum angustum* S. T. Blake, *Sorghum intrans* F. Muell. ex Benth., *Sorghum ecarinatum* Lazarides, *Sorghum bulbosum* Lazarides, *Sorghum plumosum* (R. Br.) P. Beauv., *Sorghum stipoides* (Ewart & Jean White) C. A. Gardner & C. E. Hubb., *Sorghum interjectum* Lazarides, and *Sorghum exstans* Lazarides, all of which are endemic to Australia. Among these, *S. interjectum* and *S. plumosum* are perennial species with $2n = 30, 40$ and $2n = 10, 20, 30$ respectively and the rest are annual species with $2n = 10$. Interestingly, *S. ecarinatum*, *S. bulbosum*, *S. plumosum*, *S. stipoides*, *S. interjectum*, and *S. ecarinatum* are distributed in both the Northern Territory and Western Australia, whereas *S. amplum* and *S. brachypodium* can only be found in Western Australia and the Northern Territory respectively. *Sorghum intrans* is found in north-western Northern Territory and *S. exstans* is found on Melville Island and adjoining mainland. These species have small sessile spikelet and a well-developed pedicellate spikelet (Lazarides et al., 1991).

Heterosorghum

Sorghum laxiflorum F. M. Bailey is the sole member of *Heterosorghum*, and is native to Australia and New Guinea (Price et al., 2005a). In Australia, it is commonly found in Northern Territory and Queensland. It is an annual $2n = 40$ plant with a comparatively large, sessile spikelet, obovoid to ellipsoid caryopsis and reduced spikelets (Lazarides et al., 1991).

Chaetosorghum

Sorghum macrospermum E. D. Garber is the sole member of *Chaetosorghum* and is endemic to the Northern Territory (isolated to limestone outcrops around Katherine) (Price et al., 2005a). It is an annual, $2n = 40$ species and has a small, sessile spikelet with an ovoid to ellipsoid caryopsis as well as a reduced pedicellate spikelet (Lazarides et al., 1991).

PHYLOGENETIC RELATIONSHIPS OF THE GENUS SORGHUM

The phylogenetic relationships within the genus *Sorghum* are complex, with several unresolved and potentially controversial issues. The primary gene pool (GP-1) of sorghum contains the cultivated species, *S. bicolor* and the wild species *S. propinquum* (Harlan and de Wet, 1971). The remaining members of *Eusorghum*, *S. halepense* and *S. xalmum*, belong to the small secondary gene pool (GP-2) (Stenhouse et al., 1997; Dillon et al., 2001). *Sorghum* has a comparatively larger tertiary gene pool (GP-3) which includes all the species in the other four subgenera. Members of GP-1 and GP-2 are closely related to each other whereas the members of GP-3 are more distantly

related (Harlan and de Wet, 1971). GP-3 species potentially contain many important genetic resources for sorghum improvement. However, species of this gene pool have been poorly studied as they are restricted to specific geographical areas (Bhattacharya et al., 2011).

The availability of the *S. bicolor* genome (Paterson et al., 2009) has facilitated phylogenetic studies of sorghum species based upon molecular analysis. To date, there have been several studies into the sorghum phylogeny based on nuclear genomic information together with some chloroplast genomic data. In a study by Sun et al. (1994) the *ITS* region of 13 sorghum species were sequenced covering all the subgenera, revealing the very close relationships within *Eu-sorghum*. *Sorghum bicolor* was found to be more closely related to *S. nitidum*. However, a similar study by Spangler et al. (1999), sequencing the *ndhF* gene of 39 species of the Tribe Andropogoneae, suggested some contrary relationships within the genus as a whole. For example, they suggested a distant relationship of *S. nitidum* with *S. bicolor* by being closely related to *S. laxiflorum*, and many other opposite relationships compared to the results of Sun et al. (1994) (Spangler et al., 1999).

In an attempt to clarify these contradictory classifications, Spangler (2003), presented a revised unranked classification for the genus *Sorghum* based on molecular and morphological evidence. According to this classification, *Sorghum* can be divided into three genera namely, genus *Sorghum*, genus *Vacoparis* and genus *Sarga*. Although the relationships within these three genera are still unknown, some of the changes have been already accepted by International Code of Botanical Nomenclature (ICBN). The genus *Sorghum* in this classification contains three species (*Sorghum bicolor*, *Sorghum halepense*, and *Sorghum nitidum*), the genus *Vacoparis* contains two species (*Vacoparis macrospermum* and *Vacoparis laxiflorum*), and the remaining genus *Sarga*, comprises eight species; *Sarga angustum*, *Sarga intrans*, *Sarga leiocladum*, *Sarga plumosum*, *Sarga purpureosericeum*, *Sarga timorensis*, *Sarga trichocladum*, and *Sarga versicolor*, which was created by collapsing sixteen species of sorghum into eight species.

Conversely, a more recent study (Dillon et al., 2007b), used 25 *Sorghum* taxa (not including *S. trichocladum*) to resolve the complex phylogeny of the *Sorghum* genus as many of the previous studies have resulted in contradictory classifications (Sun et al., 1994; Spangler et al., 1999; Dillon et al., 2001; Spangler, 2003; Dillon et al., 2004). Using a combined molecular analysis of *ITS1*, *ndhF*, and *Adh1*, all the sorghum species were placed in a monophyletic clade with two distinct lineages. The subgenus *Eu-sorghum* was in the same clade as *Chaetosorghum* and *Heterosorghum*, consistent with the close relationship of these two later subgenera to cultivated sorghum that was reported in earlier studies (Spangler et al., 1999; Dillon et al., 2001; Spangler, 2003; Dillon et al., 2004) and was later proved by a study by Ng'uni et al. (2010). The very close relationship of these two subgenera has been found in many other studies using morphological, cytogenetic and molecular studies despite their being considered as two separate subgenera. In addition, another clearly discrete clade was observed with all the *Parasorghum* and *Stiposorghum* species with three different clusters including *S. brachypodum* and *S. matarankense* in one cluster, *S. interjectum* and *S. ecarinatum* in another cluster,

and *S. exstans*, *S. intrans* and *S. angustum* in the third cluster. The rest of the seven species in those two subgenera formed an unresolved polytomy within this clade with no clear separation for these species. Most importantly, this study demonstrates that most of the modifications in the revised classification of Spangler (2003) are not valid except for placing *Chaetosorghum* and *Heterosorghum* together in one section. Clearly more molecular evidence is required before reclassifying the genus *Sorghum* into three subgenera.

An alternative explanation for the confusion around the *Sorghum* taxonomy is the possibility that the genus is polyphyletic within the tribe Andropogoneae (Hawkins et al., 2015). Hawkins et al. (2015) compared four nuclear loci data in 16 sorghum species together with 57 species in Andropogoneae and were able to identify two major lineages; clade I: *Eu-sorghum*, *Chaetosorghum* and *Heterosorghum*, and clade II: *Stiposorghum* and *Parasorghum* supporting previous studies done by Duvall and Doebley (1990); Sun et al. (1994), and Dillon et al. (2001). These studies were able to provide evidence of the sister relationships of these species to *Eu-sorghum* that was contrary to the single genus *Vacoparis* proposed by Spangler (2003). In clade II of the study of Hawkins et al., 2015, *S. matarankense* (*Parasorghum*) is resolved within *Stiposorghum* suggesting that it might belong to *Stiposorghum* or *Parasorghum* might be paraphyletic. However, the relationships within the clade *Stiposorghum* were only supported by low bootstrap values making them more difficult to resolve.

GENE FLOW BETWEEN WILD AND CULTIVATED SORGHUM

It has been found that many major crops are capable of natural hybridization with their wild relatives (Ellstrand et al., 1999) due to the fact that they are biologically in the same genus as their wild progenitors (Harlan and de Wet, 1971). The introgression of genes from wild relatives into crops supports the increasing genetic diversity of many species (Arnold, 2004). It is well known that diversity of wild progenitors is usually higher than that of the corresponding cultivated varieties. This is a result of domestication in which the bottleneck effect has limited the genetic diversity (Papa et al., 2005). Thus, the wild relatives of the crops may harbor valuable genetic resources and unique sources of diversity. Many studies have been carried out to study the extent and direction of the gene flow in crop-wild population complexes such as maize (Hufford et al., 2013), barley (Jakob et al., 2014), and rye (Schreiber et al., 2019), but studies on sorghum are limited. *Sorghum bicolor* subsp. *bicolor* has the advantage of having a wild progenitor, subsp. *verticilliflorum*, and its weedy relative, *S. drummondii*, which are interfertile with the cultivated species, and also grow sympatrically with cultivated forms (de Wet et al., 1970; de Wet, 1978). Studies have been done to detect the direction of gene flow through the cultivated, wild and weedy forms of sorghum, mainly based on the agricultural regions in Kenya (Mutegi et al., 2010; Mutegi et al., 2012), Ethiopia and Niger (Tesso et al., 2008), northern Cameroon (Barnaud et al., 2009), and western Africa (Sagnard

et al., 2011). All of these studies have had the same conclusion, suggesting that the crop-to-wild gene flow is more common. The studies have also emphasized a close genetic relationship between wild and crop species of sorghum.

Mutegi et al. (2012) concluded that gene flow is asymmetric by proving the rate of gene flow from crop-to-wild is higher than the gene flow from wild-to-crop, and also proposed three scenarios that could affect this asymmetric gene flow. Firstly, the sizes of the crop and wild populations might be a reason for this asymmetric gene flow which favors the larger population size of the crops compared to the smaller size of the wild populations in most agricultural lands in Africa. Farmers tend to remove wild progenitors of sorghum, considering them to be weeds. As a result, the cultivated sorghum plants produce more pollen than the wild sorghum, resulted in higher rates of pollen flow from crop to wild. Secondly, differences in the mating systems between cultivated and wild sorghum species could be a contributing factor. The higher rates of outcrossing in wild sorghums relative to cultivated sorghums facilitate the cross pollination. Thirdly, seed selection by farmers has an effect on the asymmetric gene flow. According to this concept, farmers selecting against early generations of hybrids can reduce the possibility of gene introgression from wild plants to cultivated plants. The gene flow between cultivated and wild forms has played a key role in producing intermediate species of sorghum. Doggett (1965), suggested that the balance between natural selection for wild traits and farmer selection for cultivated traits resulted in the great genetic diversity of sorghum.

CURRENT ISSUES WITH SORGHUM

Genetic Bottlenecks

The wild ancestors of sorghum have various advantageous traits such as palatable grains, high yield, wide distribution, and higher abundance over large areas. As a result, they became the main food source of early people in African savannah (de Wet and Shechter, 1977). However, with the process of domestication, most of these morphological traits were changed due to automatic selection. Tillering of the plants as well as aerial branching were reduced to have plants with only one main stem with a single inflorescence which ultimately resulted in uniform maturity. An extremely compact inflorescence was produced by contracting the axis and branches. The grain size became larger as a result of an increase in the amount of endosperm and subsequently the shape of the grain changed from elliptic to become more obovate. The breakable spikelet clusters changed to one remaining attached to the rachises at maturity (Venkateswaran et al., 2019a).

Many studies supported the concept of co-existence of wild sorghum with the cultivated sorghum in many agricultural fields of Africa (Barnaud et al., 2009; Mutegi et al., 2010; Mutegi et al., 2012). Using pure cultivated, pure wild and putative hybrids, they have proven that there is a clear genetic divergence between the populations of pure wild and pure cultivated. Interestingly, the putative hybrid group played an important role in terms of

genetic diversity by having an intermediate position in between the pure wild and pure cultivated populations. Genetic diversity reduction is known to be a result of domestication. According to the study of Mutegi et al. (2011), the genetic diversity of wild sorghum is significantly higher than the genetic diversity of cultivated sorghum in Kenya. These results agreed with the results of similar studies of Barnaud et al. (2009) and Sagnard et al. (2011) which were carried out at a local scale in Cameroon and national scale in Mali and Guinea respectively. In contrast, a parallel study carried out on a local scale by Mutegi et al. (2012) indicated that the genetic diversity between these two groups were similar in terms of gene diversity, allelic richness, and private allelic richness. However, they were able to discover 19 unique alleles in cultivated sorghum and 31 unique alleles in wild sorghums suggesting that the two gene pools were able to preserve their genetic diversity to some extent even if they were subjected to gene flow. These rare alleles of the wild plants might be linked with the traits such as drought tolerance or disease resistance.

A more recent study of Fernandez et al. (2014) has assessed the genetic diversity of landraces and wild/weedy relatives of sorghum in western Kenya using SSR markers. These authors have concluded that wild sorghum populations harbor a higher genetic diversity relative to the cultivated forms. Furthermore, in the cluster analysis although the cultivated and wild forms formed separate groups, the weedy hybrids failed to have a separate cluster from the wild forms suggesting that so called “hybrids” are closely related to the wild sorghums. Fernandez et al. (2014) outlined several reasons for this reduced gene flow and genetic diversity. For instance; farmer selection for desired traits and agronomic practices such as weeding have limited the gene flow and diversity within the cultivated species by means of reducing the cross pollination between wild and cultivated sorghums (Okeno et al., 2012). There might be several reasons for these controversial conclusions of genetic diversity differences between the wild and cultivated populations of sorghums. Differences in the experimental design, experimental area, sample size, and number of markers can affect the results of these studies. Therefore, a broader scale study which covers almost all the regions and species of sorghum is required for further validation these concepts.

Although many studies have indicated that the genetic diversity of sorghum has been reduced due to domestication, the study of Venkateswaran et al. (2019a) claimed that the variability of the plant species within the group has increased with domestication. Authors have stated that the variability in plant types, spikelet types, grain types, and inflorescence types as well as the distribution of the species have been greatly increased with the process of domestication. The morphological changes associated with domestication often gave rise to adaptations to new environments which enhanced the range of the species. These new characteristics were fixed to the new group of cultivated sorghum plants.

Grain sorghum farmers have been facing difficulty in attempts to increase yields per unit of land. While most other major cereal crops have shown significant improvement in yield gains in the

past 50 years, sorghum yields have plateaued, with levels peaking in 1981 (**Figure 2**) (Mason et al., 2008; Aruna and Cheruku, 2019). One potential cause of this plateau is a low rate of genetic enhancement, with breeding programs for other crops generally receiving more funding than sorghum during this time, and consequently being more successful (Frey, 1996; Mason et al., 2008). This is problematic due to the unlikeliness of future increases in sorghum production through a greater availability of farmland, meaning the majority of increases must come from further intensification of farming. Successful yield gains through breeding have sometimes led to losses in other crop traits, such as various Indian sorghum lines being created and used specifically for high yields despite reductions in grain quality (Aruna and Cheruku, 2019). We must continue to tackle sorghum's genetic homogeneity issues, increasing the amount of research done and the breadth of methods used, in order to increase yields again without sacrificing nutrition.

CROP WILD RELATIVES IN SORGHUM IMPROVEMENT

Barriers to Use of Sorghum's Wild Relatives

Undomesticated sorghum species harbor beneficial traits which can be employed as prospective markers to the phylogenetic relationships within the genus as well as between similar plant families. One of the major constraints to utilizing these genetic resources of wild relatives is the barriers to gene transfer between cultivated crops and their wild relatives (Bevan et al., 2017). Some of the sorghum species in the primary and secondary gene pools have been extensively used in genetic studies since they have few genetic incompatibilities with *S. bicolor*. However, most of the wild sorghum species belong to the tertiary gene pool and gene transfer to the cultivated sorghum species is difficult. Recent phylogenetic studies have revealed the two undomesticated species *S. laxiflorum* and *S. macrospermum* as the most closely related species to the cultivated sorghum species (Dillon et al., 2007c). Many unsuccessful attempts have been made to produce viable hybrids (Garber and Snyder, 1951; Sun et al., 1994; Huelgas et al., 1996). Gene transfer from the crop wild relatives to the

cultivated sorghum species is challenging for several reasons. The main reason is the strong pre- and post-zygotic reproductive barriers between wild and domesticated species. These sterility barriers can be seen as a result of differences in genome size, chromosome morphology, pollen-pistil incompatibilities, and embryo abortions (Garber and Snyder, 1951; Price et al., 2005b). Hybrid embryo formation may be impossible due to the pollen-pistil incompatibilities between *S. bicolor* and wild species (Hodnett et al., 2005). However, successful efforts of hybridization have been reported with artificial hybridization techniques such as embryo rescue (Price et al., 2005b). One successful attempt has been reported between the species *S. bicolor* and *S. macrospermum*, using embryo rescue methods (Price et al., 2005b). Techniques such as the use of bridge species, irradiation of pollen grains, and chromosome doubling have also been used to overcome these sterility barriers (Kumari et al., 2016). Kuhlman et al. (2010) have also successfully developed a *S. bicolor* line which is homozygous for the recessive *iap* (inhibition of alien pollen) gene, allowing pollen tubes to grow to completion, even when the pollen is from a GP-3 species. Hybrids have since been made by crossing *S. bicolor* with *S. macrospermum* (Kuhlman et al., 2010), and also with *Saccharum* spp. (Hodnett et al., 2010). A detailed account of attempts of producing hybrids between cultivated sorghum and wild sorghum has been explained in a review of Ohadi et al. (2017) (**Table 2**).

Additionally, sorghum CWR may have been overlooked historically due to their apparent lack of usefulness regarding advantageous agricultural traits. This trend has been seen in many CWR (Jansky et al., 2013), including wild sorghum species having been overlooked in the past due to their low yields and “weedy” characteristics (Cox et al., 1984; McWhorter, 1989). However, there are various reasons why *Sorghum* species should no longer be viewed this way. Several of these species have been shown to possess traits which would be desirable in sorghum crops (Kamala et al., 2002; Venkateswaran, 2003; Cowan et al., 2020). Increasing genetic heterogeneity through hybridization can also be unexpectedly beneficial through heterosis—enhancement of traits through mixing genes of two genetically-distinct parents. Some benefits can be phenotypically obvious, for example with Jordan et al. (2004) finding some hybrids of *S. bicolor* subsp. *bicolor* and *verticilliflorum* with higher yields than either of the parent plants. This finding was surprising given that subsp. *verticilliflorum* typically has low grain yields. Other benefits of heterosis might be less immediately noticeable, including reduced susceptibility to pests, pathogens, and environmental changes (Chen, 2010).

Use of Gene Pools 1 and 2

Due to the incompatibility of crossing *S. bicolor* with species in GP-3, most existing hybrids have been made through crosses of *S. bicolor* with members of gene pools 1 and 2 (Duncan et al., 1991). These include: *S. bicolor* subsp. *verticilliflorum* (Cox et al., 1984; Jordan et al., 2004) and *S. propinquum* (Wooten, 2001) being used to increase yield; *S. halepense* being used to introduce perennialism (Cox et al., 2002; Dweikat, 2005); and *S. propinquum* being used to increase height and earliness of development (Wooten, 2001) (**Table 3**). There have also been countless crosses between

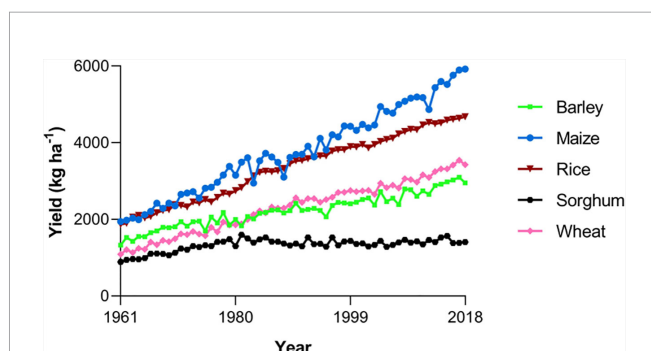


FIGURE 2 | Trends in the total yields of the world's five most important cereal crops. Data obtained from FAO (2019).

TABLE 2 | Experimental details of hybridization between *S. bicolor* and its wild relatives (Ohadi et al., 2017).

Taxon	Status	References
<i>S. bicolor</i> and <i>S. alnum</i>	Successful hybrids	Endrizzi, 1957
<i>S. bicolor</i> and <i>S. angustum</i>	Unsuccessful (<i>in vivo</i> rescue of the developing embryos were required)	Price et al., 2006
<i>S. bicolor</i> and <i>S. bicolor</i> nothosubsp. <i>drummondii</i>	Unassisted hybridization	Schmidt et al., 2013
<i>S. bicolor</i> and <i>S. bicolor</i> nothosubsp. <i>drummondii</i>	Successful hybrids	Werle et al., 2014
<i>S. bicolor</i> and <i>S. halepense</i>	Successful hybrids	Endrizzi, 1957; Hadley, 1958; Sangduen and Hanna, 1984; Piper and Kulakow, 1994; Cox et al., 2002; Dweikat, 2005; Magomere et al., 2015
<i>S. bicolor</i> and <i>S. halepense</i>	Natural introgression	Morrell et al., 2005
<i>S. bicolor</i> and <i>S. macrospermum</i>	Successful introgression using embryo rescue	Price et al., 2006; Kuhlman et al., 2010
<i>S. bicolor</i> and <i>S. nitidum</i>	Unsuccessful (<i>in vivo</i> rescue of the developing embryos were required)	Price et al., 2006
<i>S. bicolor</i> and <i>S. propinquum</i>	Successful hybrids but no use in sorghum improvement	Paterson et al., 1995; Wooten, 2001
<i>S. bicolor</i> and <i>S. versicolor</i>	Successful hybrids	Sun et al., 1991
<i>S. bicolor</i> and <i>S. bicolor</i> subsp. <i>verticilliflorum</i>	Successful hybrids	Cox et al., 1984; Jordan et al., 2004

different commercial lines of the crop (Rosenow and Clark, 1982; Duncan et al., 1991; Aruna and Cheruku, 2019).

There are also various other traits within gene pools 1 and 2 which have been listed as potentially useful for introgression into *S. bicolor*. Harlan (1992), reported that the wild race *arundinaceum* was adapted to growing in wet climates, an adaptation not common in cultivated sorghum species. The wild race *virgatum* can grow in drought conditions and their seeds have been shown to be tolerant to high temperatures (Bramel-Cox and Cox, 1988). In addition, Bramel-Cox and Cox (1988), showed that high yielding wild species *arundinaceum*, *virgatum*, and *verticilliflorum*, could be used to increase the yield of domesticated sorghum. These wild races also have resistance to the parasitic weed *Striga asiatica* Lour., a useful trait in sorghum cultivation (Rich et al., 2004). Other potentially useful traits in sorghum's GP-1 and GP-2 include: *S. bicolor* nothosubsp. *drummondii*'s allelopathic properties, which reduce the growth of weeds in the cultivated field (Baerson et al., 2008), and resistance to ergot (Tsukiboshi et al., 1998) and nematodes (Mojtahedi et al., 1993; Viaene and Abawi, 1998); and *S. halepense*'s resistance to pests such as green bug, chinch bug, and sorghum shoot fly (Nwanze et al., 1995; Dweikat, 2005) (Table 3). Meanwhile, continued crosses between commercial lines will continue to contribute to recombination efforts, while also potentially generating serendipitous new phenotypes (such as yield gains) through heterosis.

Use of Gene Pool 3

Although sorghum's GP-3 has not yet extensively been used in crop improvement, it potentially contains a high level of genetic diversity for use in sorghum improvement. This diversity is suggested by the ability of these species to adapt to a range of edaphic conditions, with Australia's native sorghums collectively covering diverse habitats including rocky slopes, sand dunes, grasslands, and forests (Lazarides et al., 1991). The niche diversity of GP-3 is much greater than that of GP-1 and GP-2, potentially providing genetic resources with which the environmental tolerances of sorghum crops could be expanded. For example, there has been great interest in increasing sorghum's tolerance to cold temperatures in order to greatly expand the zone in which it can be grown (Fiedler et al., 2016; Yu and Tuinstra, 2001). *Sorghum leiocladum* could be a good candidate species for cold tolerance genes due to its presence in temperate regions of New South Wales and Victoria, Australia. Similarly, Cowan et al., 2020, found multiple GP-3 species with greater tolerance to drought than domesticated sorghum, including *S. brachypodium* and *S. macrospermum*. Further research into GP-3 could unveil more environments to which wild species could offer novel tolerance genes. Species across GP-3 have also been shown to be resistant to biotic stressors including sorghum shoot fly (Venkateswaran, 2003; Kamala et al., 2009), spotted stem borer (Venkateswaran, 2003), and downy mildew (Kamala et al., 2002), as well as *S. angustum*, *S. amplum* and *S. bulbosum* all showing resistance to egg laying by sorghum midge (Sharma and Franzmann, 2001) (Table 3). The identification of such traits despite the limited number of studies conducted on sorghum's GP-3 suggests that there is high potential for finding further agronomically advantageous traits in this gene pool.

Cowan et al. (2020) also found that, in contrast to other cyanogenic crops, the leaf cyanogenic glucoside content of drought stressed wild sorghums is lower than that of the cultivated species (Cowan et al., 2020). Interestingly, findings of this study revealed that drought stress significantly increased the dhurrin concentration of the aboveground parts of *S. bicolor*, while the wild species were not significantly affected. Specifically, the two wild species *S. macrospermum* and *S. brachypodium* were able to maintain a higher growth rate and an insignificant aboveground dhurrin content. The regulation of the formation cyanogenic glucosides in wild sorghum species has not yet been studied in detail or compared to that of cultivated *S. bicolor*. Understanding the gene expression and regulation of cyanogenesis related genes in wild relatives of sorghum would be a crucial step in utilizing the useful traits in wild sorghum in crop improvement (Cowan et al., 2020).

Priorities for Future Work

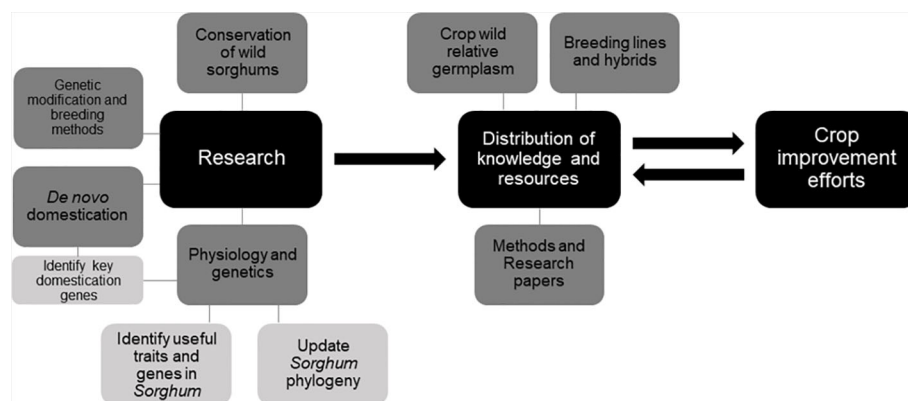
In order to maximize the impact of sorghum improvement using CWR, various steps must be taken to improve how current work is executed. These steps include further development and distribution of *S. bicolor* lines which can interbreed with species outside GP-1 and GP-2, further improvements in the sorghum GM process, increased accessibility for crop developers and researchers to CWR germplasm, knowledge, and introgression technology, and a better understanding of how

TABLE 3 | Details of potential wild sorghum species which can be used to improve cultivated sorghum.

Taxon	Gene pool	Traits	Status	References
<i>S. propinquum</i>	1	Increase grain yield, increase height, and earliness of development	Successfully introgressed to <i>S. bicolor</i>	Wooten, 2001
<i>S. bicolor</i> subsp. <i>verticilliflorum</i>	1	Increase grain yield	Successfully introgressed to <i>S. bicolor</i>	Cox et al., 1984; Jordan et al., 2004
<i>S. halepense</i>	2	Perennialism	Successfully introgressed to <i>S. bicolor</i>	Cox et al., 2002; Dweikat, 2005
<i>S. bicolor</i> subsp. <i>verticilliflorum</i>	1	Ability to grow in drought conditions, seeds with tolerance to high temperatures, high yield, parasite resistance	Reported as potential candidates for sorghum improvement	Bramel-Cox and Cox, 1988; Rich et al., 2004
<i>S. bicolor</i> nothosubsp. <i>drummondii</i>	1	Allelopathic properties, resistance to ergot and nematodes	Reported as potential candidates for sorghum improvement	Mojtahedi et al., 1993; Tsukiboshi et al., 1998; Viaene and Abawi, 1998; Baerson et al., 2008
<i>S. halepense</i>	2	Resistance to green bug, chinch bug, and sorghum shoot fly	Reported as potential candidates for sorghum improvement	Nwanze et al., 1995; Dweikat, 2005
<i>S. angustum</i>	3	Resistance to egg laying by sorghum midge	Reported as potential candidates for sorghum improvement	Sharma and Franzmann, 2001
<i>S. amplum</i>	3	Resistance to egg laying by sorghum midge	Reported as potential candidates for sorghum improvement	Sharma and Franzmann, 2001
<i>S. bulbosum</i>	3	Resistance to egg laying by sorghum midge	Reported as potential candidates for sorghum improvement	Sharma and Franzmann, 2001
<i>S. macrospermum</i>	3	Insect and disease resistance, higher growth rate and an insignificant aboveground dhurrin content under drought conditions	Successfully introgressed to <i>S. bicolor</i>	Kuhlman et al., 2008; Cowan et al., 2020
<i>S. brachypodium</i>	3	Higher growth rate and an insignificant aboveground dhurrin content under drought conditions	Reported as potential candidates for sorghum improvement	Cowan et al., 2020
<i>S. exstans</i>	3	Resistance to shoot fly	Reported as potential candidates for sorghum improvement	Kamala et al., 2009
<i>S. stipoideum</i>	3	Resistance to shoot fly	Reported as potential candidates for sorghum improvement	Kamala et al., 2009
<i>S. matarankense</i>	3	Resistance to shoot fly	Reported as potential candidates for sorghum improvement	Kamala et al., 2009
<i>S. leiocladum</i>	3	Cold tolerance	Reported as potential candidates for sorghum improvement	Fiedler et al., 2016

each of sorghum's CWR might be valuable to the crop improvement process. *De novo* domestication of wild sorghums might also be a valuable method through which new sorghum lines could be developed (Fernie and Yan, 2019). Continued research into the morphology and physiology of CWR species will allow us to determine which species are

potentially the most suitable as genetic sources for crop improvement, as well as for undergoing *de novo* domestication, taking into account potential uses, yields, and crop safety (e.g. storage of cyanogenic glucosides). Some of sorghum's key domestication genes have already been identified (Meyer and Purugganan, 2013; Tao et al., 2017). Further elucidation of the

**FIGURE 3** | Roadmap towards the use of sorghum's wild relatives in crop improvement.

Sorghum phylogeny might also help in the introgression process through allowing a better understanding of the relatedness of each species to the target crop (**Figure 3**).

All of these steps also rely on the continued conservation of the CWR species and the intraspecific genetic diversity within them. A combination of *ex situ* and *in situ* conservation techniques is vital for preserving the maximum genetic diversity (Engels et al., 2008). Currently, the world sorghum germplasm collection contains more than 200,000 accessions (FAO, 2009). Among these germplasm collections, ICRISAT (International Crops Research Institute for the Semi-Arid Tropics) has the world's depository of sorghum germplasm collection, including many accessions from GP-1 and GP-2, as well as some GP-3 accessions (Wang et al., 2015). The main GP-3 germplasm collections are located in Australia at the Australian Grains Genebank (Bhattacharya et al., 2011; Genesys-pgr, 2020) with additional germplasm—mainly of the same lines as those held by the Australian Grains Genebank—held overseas by organizations such as the USDA Agricultural Research Service and the Millennium Seed Bank. Because most *Sorghum* species are native to Australia, *in situ* protections in the nation are vital for protecting the genus' diversity. However, *in situ* protections of GP-3 species across Africa and Asia are also necessary, as these represent the genetic resources which are most easily crossed with the crop.

CONCLUSION

Sorghum is an immensely valuable multipurpose crop with several end user products. The genus *Sorghum* is rich in diversity with a highly beneficial reservoir of untapped genetic resources, especially in the tertiary gene pool. The wild relatives of sorghum contain many expedient traits which can be utilized in crop improvement. However, exploitation of these extremely valuable traits in crop improvement is still hindered due to the

limited availability of genetic information on these wild sorghum species. Furthermore, the genetic barriers in gene transfer between wild and cultivated sorghum species are challenging. However, with recent advances in next generation sequencing technologies, more genomic data will become available to researchers. This will extend the development of sorghum improvement programs using the rich, yet unexploited genetic resources in sorghum's wild relatives. These resources may also support the development of new crops in the tribe Andropogoneae (Dillon et al., 2007c) either by introgression of useful genes into genera such as *Saccharum* or by domestication of further *Sorghum* species.

AUTHOR CONTRIBUTIONS

GA and HM led the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by an Australian Research Council Discovery Project grant to RG and RH (DP180101011).

ACKNOWLEDGMENTS

Authors would like to acknowledge the contribution of Dr. Cecilia Blomstedt, School of Biological Sciences, Monash University for providing comments on the manuscript. HM was supported by a Monash University Research Training Stipend and a Tim Healey Memorial Scholarship from the AW Howard Memorial Trust. GA was supported by University of Queensland Research Training Stipend.

REFERENCES

- Anzoua, K. G., Yamada, T., and Henry, R. J. (2011). "Miscanthus," in *Wild Crop Relatives: Genomic and Breeding Resources: Industrial Crops*. Ed. C. Kole (Berlin, Heidelberg: Springer Berlin Heidelberg), 157–164.
- Appa, R., Prasada Rao, K., Mengesha, M., and Reddy, V. (1996). Geographical distribution, diversity and gap analysis of east african sorghum collection conserved at the ICRISAT Genebank. *Genet. Resour. Crop Ev.* 43 (6), 559–567. doi: 10.21475/ajcs.17.11.04.pne330
- Arnold, M. L. (2004). Natural hybridization and the evolution of domesticated, pest and disease organisms. *Mol. Ecol.* 13, 997–1007. doi: 10.1111/j.1365-294X.2004.02145.x
- Aruna, C., and Cheruku, D. (2019). "Genetic improvement of grain sorghum," in *Breeding Sorghum for Diverse End Uses*. Eds C. Aruna, K.B.R.S. Visarada, B.V. Bhat and V. A. Tonapi (Cambridge, United Kingdom: Woodhead Publishing), 157–173.
- Baerson, S. R., Dayan, F. E., Rimando, A. M., Nanayakkara, N. P. D., Liu, C., Schröder, J., et al. (2008). Functional genomics investigation of allelochemical biosynthesis in *Sorghum bicolor* root hairs. *J. Biol. Chem.* 283, 3231–3247. doi: 10.1074/jbc.M706587200
- Barnaud, A., Deu, M., Garine, E., Chantreau, J., Bolteu, J., Koïda, E. O., et al. (2009). A weed–crop complex in sorghum: The dynamics of genetic diversity in a traditional farming system. *Am. J. Bot.* 96, 1869–1879. doi: 10.3732/ajb.0800284
- Bevan, M. W., Uauy, C., Wulff, B. B. H., Zhou, J., Krasileva, K., and Clark, M. D. (2017). Genomic innovation for crop improvement. *Nature* 543, 346–354. doi: 10.1038/nature22011
- Bhattacharya, A., Rice, N., Shapter, F. M., Norton, S. L., and Henry, R. J. (2011). "Sorghum," in *Wild Crop Relatives: Genomic and Breeding Resources: Cereals*. Ed. C. Kole (Berlin, Heidelberg: Springer Berlin Heidelberg). doi: 10.1007/978-3-642-14228-4
- Bonnett, G. D., and Henry, R. J. (2011). "Saccharum," in *Wild Crop Relatives: Genomic and Breeding Resources: Industrial Crops*. Ed. C. Kole (Berlin, Heidelberg: Springer Berlin Heidelberg), 165–177.
- Bramel-Cox, P. J., and Cox, T. S. (1988). "Use of wild sorghums in sorghum improvement," in *Proceedings of the 43rd Annual Corn and Sorghum Industry Research Conference, USA*. Ed. W. D. (Washington DC: American Seed Trade Association), 13–26.
- Brozynska, M., Furtado, A., and Henry, R. J. (2016). Genomics of crop wild relatives: expanding the gene pool for crop improvement. *Plant Biotechnol. J.* 14 (4), 1070–1085. doi: 10.1111/pbi.12454
- Chen, Z. J. (2010). Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* 15, 57–71. doi: 10.1016/j.tplants.2009.12.003
- Cowan, M. F., Blomstedt, C. K., Norton, S. L., Henry, R. J., Möller, B. L., and Gleadow, R. (2020). Crop wild relatives as a genetic resource for generating low-cyanide, drought-tolerant Sorghum. *Environ. Exp. Bot.* 169:103884. doi: 10.1016/j.envexpbot.2019.103884

- Cox, T., House, L., and Frey, K. J. (1984). Potential of wild germplasm for increasing yield of grain sorghum. *Euphytica* 33 (3), 673–684. doi: 10.1007/BF00021895
- Cox, T., Bender, M., Picone, C., Tassel, D. V., Holland, J., Brummer, E., et al. (2002). Breeding perennial grain crops. *CRC Crit. Rev. Plant Sci.* 21 (2), 59–91. doi: 10.1080/0735-260291044188
- Dahlberg, J. A. (2000). “Classification and characterization of sorghum,” in *Sorghum*. Eds. C. W. Smith and J. R. Frederick (New York, USA: John Wiley and Sons).
- De Wet, J. M. J., and Harlan, J. R. (1971). The origin and domestication of *Sorghum bicolor*. *Econ. Bot.* 25, 128–135. doi: 10.1007/BF02860074
- De Wet, J., and Shechter, Y. (1977). “Evolutionary dynamics of sorghum domestication,” in *Crop Resources*. Ed. D. S. Seigler (New York, USA: Academic Press), 179–191. doi: 10.1016/B978-0-12-634950-4.50020-2
- De Wet, J. M. J., Harlan, J. R., and Price, E. G. (1970). Origin of variability in the Spontanea complex of *Sorghum bicolor*. *Am. J. Bot.* 57, 704–707. doi: 10.2307/2441294
- De Wet, J. M. J. (1978). Systematics and evolution of *Sorghum* Sect. *Sorghum* (Gramineae). *Am. J. Bot.* 65 477–484. doi: 10.1002/j.1537-2197.1978.tb06096.x
- Dempewolf, H., Eastwood, R. J., Guarino, L., Khoury, C. K., Müller, J. V., and Toll, J. (2014). Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecol. Sust. Food.* 38, 369–377. doi: 10.1080/21683565.2013.870629
- Dillon, S. L., Lawrence, P. K., and Henry, R. J. (2001). The use of ribosomal ITS to determine phylogenetic relationships within *Sorghum*. *Plant Syst. Evol.* 230, 97–110. doi: 10.1007/s006060170007
- Dillon, S. L., Lawrence, P. K., Henry, R. J., Ross, L., Price, H. J., and Johnston, J. S. (2004). *Sorghum laxiflorum* and *S. macrospermum*, the Australian native species most closely related to the cultivated *S. bicolor* based on ITS1 and ndhF sequence analysis of 25 *Sorghum* species. *Plant Syst. Evol.* 249, 233–246. doi: 10.1007/s00606-004-0210-7
- Dillon, S. L., Lawrence, P. K., and Henry, R. J. (2007a). The new use of *Sorghum bicolor*-derived SSR markers to evaluate genetic diversity in 17 Australian *Sorghum* species. *Plant Genet. Resour.* 3, 19–28. doi: 10.1079/PGR200454
- Dillon, S. L., Lawrence, P. K., Henry, R. J., and Price, H. J. (2007b). *Sorghum* resolved as a distinct genus based on combined ITS1, ndh F and Adh 1 analyses. *Plant Syst. Evol.* 268, 29–43. doi: 10.1007/s00606-007-0571-9
- Dillon, S. L., Shapter, F. M., Henry, R. J., Cordeiro, G., Izquierdo, L., and Lee, L. S. (2007c). Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). *Ann. Bot.* 100, 975–989. doi: 10.1093/aob/mcm192
- Doggett, H. (1965). Disruptive selection in crop development. *Nature* 206, 279–280. doi: 10.1038/hdy.1968.1
- Doggett, H. (1988). *Sorghum* (Harlow, Essex, England: Longman Scientific & Technical New York: Wiley).
- Duncan, R., Bramel-Cox, P., and Miller, F. (1991). “Contributions of introduced sorghum germplasm to hybrid development in the USA,” in *Use of Plant Introductions in Cultivar Development Part 1*. Eds. H. L. Shands and L. E. Wiesner (Wisconsin, USA: CSSA), 69–102.
- Duvall, M. R., and Doebley, J. F. (1990). Restriction Site Variation in the chloroplast genome of *Sorghum* (Poaceae). *Syst. Bot.* 15, 472–480. doi: 10.2307/2419363
- Dweikat, I. (2005). A diploid, interspecific, fertile hybrid from cultivated sorghum, *Sorghum bicolor* and the common Johnsongrass Weed *Sorghum halepense*. *New Strateg. Plant Improvement.* 16, 93–101. doi: 10.1007/s11032-005-5021-1
- Ehret, C. (2014). “Linguistic evidence on the origins of food production in Africa: where are we now?”, in *Archaeology of African Plant Use*. Eds. C. J. Stevens, S. Nixon and M. A. Murray (California, USA: Left Coast Press, Inc.), 233–242.
- Ellstrand, N. C., Prentice, H. C., and Hancock, J. F. (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Annu. Rev. Ecol. Syst.* 30, 539–563. doi: 10.1146/annurev.ecolsys.30.1.539
- Endrizzi, J. E. (1957). Cytological studies of some species and hybrids in the Eu-sorghums. *Bot. Gaz.* 119, 1–10. doi: 10.1086/335954
- Engels, J., Maggioni, L., Maxted, N., and Dulloo, M. (2008). “Complementing in situ conservation with ex situ measures,” in *Conserving plant genetic diversity in protected areas: Population management of crop wild relatives* (Rome: CAB Int.), 169–181.
- FAO (2009). *Commission on Genetic Resources for Food and Agriculture* (Rome). Available at: <http://www.fao.org/3/a-k6276e.pdf> (Accessed 28th December 2019).
- FAO (2019). *FAOSTAT Statistical Database*. Available at: <http://www.fao.org/faostat/> (Accessed 28th December 2019).
- FAOSTAT (2019). *Production/Yield quantities of Sorghum in World* (Rome). Available at: <http://www.fao.org/faostat/en/#data/QC/visualize> (Accessed 28th December 2019).
- Fernandez, M. G. S., Okeno, J. A., Mutege, E., Fessehaie, A., and Chalfant, S. (2014). Assessment of genetic diversity among sorghum landraces and their wild/weedy relatives in western Kenya using simple sequence repeat (SSR) markers. *Conserv. Genet.* 15, 1269–1280. doi: 10.1007/s10592-014-0616-x
- Fernie, A. R., and Yan, J. (2019). De novo domestication: An alternative route toward new crops for the future. *Mol. Plant* 12 (5), 615–631. doi: 10.1016/j.molp.2019.03.016
- Fiedler, K., Bekele, W. A., Matschegewski, C., Snowdon, R., Wieckhorst, S., Zacharias, A., et al. (2016). Cold tolerance during juvenile development in sorghum: a comparative analysis by genomewide association and linkage mapping. *Plant Breed.* 135 (5), 598–606. doi: 10.1111/pbr.12394
- Frey, K. J. (1996). *National plant breeding study-1: Human and financial resources devoted to plant breeding research and development in the United States in 1994*. (Iowa, USA: Iowa State University), 98, 51.
- Fuller, D. Q., and Stevens, C. J. (2018). “Sorghum domestication and diversification: a current archaeobotanical perspective,” in *Plants and People in the African Past*. Eds. A. M. Mercuri, A. C. D’andrea, R. Fornaciari and A. Höhn (Berlin/Heidelberg, Germany: Springer), 427–452. doi: 10.1007/978-3-319-89839-1_19
- Fuller, D. Q. (2003). “African crops in prehistoric South Asia: a critical review,” in *Food, fuel and fields: Progress in African archaeobotany*. Eds. K. Neumann, A. Butler and S. Kahlheber (Köln, Germany: Heinrich-Barth-Institute), pp. 239–271.
- Garber, E. D., and Snyder, L. A. (1951). Cytotaxonomic studies in the genus *Sorghum*. II. Two new species from Australia. *Madroño* 11, 6–10.
- Genesys-pgr (2020). Available at: <https://www.genesys-pgr.org> (Accessed 29th April 2020).
- Gupta, S. C., de Wet, M. J., and Harlan, J. R. (1978). Morphology of *Saccharum-Sorghum* hybrid derivatives. *Am. J. Bot.* 65, 936–942. doi: 10.1002/j.1537-2197.1978.tb06158.x
- Hadley, H. H. (1958). Chromosome numbers, fertility and rhizome expression of hybrids between grain sorghum and johnsongrass. *Agron. J.* 50, 278–282. doi: 10.2134/agronj1958.00021962005000050015x
- Hagerty, M. J. (1941). Comments on writings concerning Chinese sorghums. *Harv. J. Asiat. Stud.* 5, 234–260. doi: 10.2307/2717914
- Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* 156 (1-2), 1–3. doi: 10.1007/s10681-007-9363-0
- Hariprasanna, K., and Patil, J. V. (2015). “Sorghum: Origin, classification, biology and improvement,” in *Sorghum molecular breeding*. Eds. R. Madhusudhana, P. Rajendrakumar and J. V. Patil (New Delhi, India: Springer), 3–20. doi: 10.1007/978-81-322-2422-8_1
- Harlan, J. R., and De Wet, J. M. J. (1971). Toward a rational classification of cultivated sorghums. *Crop Sci.* 12, 172–176. doi: 10.2307/1218252
- Harlan, J. R., and de Wet, J. M. J. (1972). A simplified classification of cultivated *Sorghum*. *Crop Sci.* 12, 172. doi: 10.2135/cropsci1972.0011183X001200020005x
- Harlan, J. R. (1992). “Indigenous African agriculture,” in *Crops Man*. Ed. J. R. Harlan (Wisconsin, USA: American Society of Agronomy) doi: 10.2135/1992.cropsandman.c9
- Harlan, J. R. (1995). *The living fields: Our Agricultural Heritage* (Cambridge: Cambridge University Press).
- Hawkins, J. S., Ramachandran, D., Henderson, A., Freeman, J., Carlise, M., Harris, A., et al. (2015). Phylogenetic reconstruction using four low-copy nuclear loci strongly supports a polyphyletic origin of the genus *Sorghum*. *Ann. Bot.* 116, 291–299. doi: 10.1093/aob/mcv097
- Hodnett, G. L., Burson, B. L., Rooney, W. L., Dillon, S. L., and Price, H. J. (2005). Pollen–pistil interactions result in reproductive isolation between *Sorghum bicolor* and divergent *Sorghum* species. *Crop Sci.* 45, 1403–1409. doi: 10.2135/cropsci.2004.0429
- Hodnett, G. L., Hale, A. L., Packer, D. J., Stelly, D. M., Da Silva, J., and Rooney, W. L. (2010). Elimination of a reproductive barrier facilitates intergeneric

- hybridization of *Sorghum bicolor* and *Saccharum*. *Crop Sci.* 50 (4), 1188–1195. doi: 10.2135/cropsci2009.09.0486
- Hoyt, E. (1988). *Conserving the wild relatives of crops*. (Rome: International Board for Plant Genetic Resources).
- Huelgas, V. C., Lawrence, P., Adkins, S. W., Mufti, M. U., and Goodwin, I. D. (1996). "Utilisation of the Australian native species for sorghum improvement," in *Proceedings of the 3rd Australian Sorghum Conference*. Eds. M. A. Foale, R. G. Henzell and J. F. Kniepp (Melbourne, Australia: Australian Institute of Agricultural Science), 369–375.
- Hufford, M. B., Lubinsky, P., Pyhajarvi, T., Devengenz, M. T., Ellstrand, N. C., and Ross-Ibarra, J. (2013). The genomic signature of crop-wild introgression in Maize. *PLoS Genet.* 9 (9), e1003477. doi: 10.1371/journal.pgen.1003477
- Jakob, S. S., Rodder, D., Engler, J. O., Shaaf, S., Ozkan, H., Blattner, F. R., et al. (2014). Evolutionary history of wild Barley (*Hordeum vulgare* subsp. *spontaneum*) analyzed using multilocus sequence data and paleodistribution modeling. *Genome Biol. Evol.* 6 (3), 685–702. doi: 10.1093/gbe/evu047
- Jansky, S., Dempewolf, H., Camadro, E. L., Simon, R., Zimnoch-Guzowska, E., Bisognin, D., et al. (2013). A case for crop wild relative preservation and use in potato. *Crop Sci.* 53 (3), 746–754. doi: 10.2135/cropsci2012.11.0627
- Jarvis, A., Lane, A., and Hijmans, R. (2008). The effect of climate change on crop wild relatives. *Agric. Ecosyst. Environ.* 126 (1–2), 13–23. doi: 10.1016/j.agee.2008.01.013
- Jordan, J., Butler, D., Henzell, B., Drenth, J., and McIntyre, L. (2004). *Diversification of Australian sorghum using wild relatives. Paper presented at the New Directions for a Diverse Planet* (Brisbane, Australia: Proceedings of the 4th International Crop Science Congress).
- Kamala, V., Singh, S. D., Bramel, P. J., and Rao, D. M. (2002). Sources of Resistance to Downy Mildew in Wild and Weedy Sorghums. *Crop Sci.* 42, 1357–1360. doi: 10.2135/cropsci2002.1357
- Kamala, V., Sharma, H., Manohar Rao, D., Varaprasad, K., and Bramel, P. (2009). Wild relatives of sorghum as sources of resistance to sorghum shoot fly, *Atherigona soccata*. *Plant Breed.* 128 (2), 137–142. doi: 10.1111/j.1439-0523.2008.01585.x
- Kimber, C. T. (2000). "Origins of domesticated sorghum and its early diffusion to India and China," in *Sorghum: Origin, history, technology, and production*. Eds. C. W. Smith and R. A. Frederiksen (New York, USA: John Wiley), 3–98.
- Kleih, U., Ravi, S. B., Rao, B. D., and Yoganand, B. (2000). *Industrial utilization of sorghum in India*, (India: International Crops Research Institute for the Semi-Arid Tropics).
- Kuhlman, L. C., Burson, B. L., Klein, P. E., Klein, R. R., Stelly, D. M., Price, H. J., et al. (2008). Genetic recombination in *Sorghum bicolor* × *S. macrospermum* interspecific hybrids. *Genome* 51, 749–756. doi: 10.1139/G08-061
- Kuhlman, L. C., Burson, B. L., Stelly, D. M., Klein, P. E., Klein, R. R., Price, H., et al. (2010). Early-generation germplasm introgression from *Sorghum macrospermum* into sorghum (*S. bicolor*). *Genome* 53 (6), 419–429. doi: 10.1139/g10-027
- Kumari, P., Pahuja, S. K., Arya, S., and Patil, J. V. (2016). "Sorghum," in *Broadening the genetic base of grain cereals*. Eds. M. Singh and S. Kumar (New Delhi, India: Springer), pp. 163–203.
- Lazarides, M., Hacker, J. B., and Andrew, M. H. (1991). Taxonomy, cytology and ecology of indigenous Australian sorghums (*Sorghum* Moench: *Andropogoneae*: *Poaceae*). *Aust. Syst. Bot.* 4, 591–635. doi: 10.1071/sb9910591
- Leitch, I. J., Johnston, E., Pellicer, J., Hidalgo, O., and Bennett, M. D. (2019). *Angiosperm DNA C-values database (release 9.0, Apr 2019)*. Available at: <https://cvalues.science.kew.org/> (Accessed 23rd March 2020).
- Mace, E. S., Rami, J. F., Bouchet, S., Klein, P. E., Klein, R. R., Kilian, A., et al. (2009). A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. (Research article)(Report). *BMC Plant Biol.* 9, 13. doi: 10.1186/1471-2229-9-13
- Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4, 1–9. doi: 10.1038/ncomms3320
- Magomere, T. O., Obukosia, S. D., Shibairo, S.II, Ngugi, E. K., and Mutitu, E. (2015). Evaluation of relative competitive ability and fitness of *Sorghum bicolor* × *Sorghum halepense* and *Sorghum bicolor* × *Sorghum sudanense* F1 hybrids. *J. Biol. Sci.* 15, 1–15. doi: 10.3923/jbs.2015.1.15
- Mann, J. A., Kimber, C. T., and Miller, F. R. (1983). *The Origin and Early Cultivation of Sorghum in Africa* (Texas, USA: Texas Agricultural Experiment Station).
- Martin, J. H. (1936). "Sorghum improvement," in *The Yearbook of Agriculture* (1936). Ed. United States Department of Agriculture (Washington DC, USA: USDA).
- Mason, S. C., Kathol, D., Eskridge, K. M., and Galusha, T. D. (2008). Yield increase has been more rapid for maize than for grain sorghum. *Crop Sci.* 48 (4), 1560–1568. doi: 10.2135/cropsci2007.09.0529
- McWhorter, C. G. (1989). History, biology, and control of johnsongrass. *Reviews of weed science* (USA) 4, 85–121.
- Meadow, R. H. (1996). "The origins and spread of agriculture and pastoralism in northwestern South Asia", in *The origins and spread of agriculture and pastoralism in Eurasia*. Ed. D.R. Harris (Washington DC, USA: Smithsonian), 390–412.
- Meilleur, B. A., and Hodgkin, T. (2004). In situ conservation of crop wild relatives: status and trends. *Biodivers. Conserv.* 13 (4), 663–684. doi: 10.1023/B: BIOC.0000011719.03230.17
- Meyer, R. S., and Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* 14, 840. doi: 10.1038/nrg3605
- Mojtahedi, H., Santo, G. S., and Ingham, R. E. (1993). Suppression of Meloidogyne chitwoodi with Sudangrass cultivars as green manure. *J. Nematol.* 25, 303–311.
- Morrell, P., Williams-Coplin, T., Lattu, A., Bowers, J., Chandler, J., and Paterson, A. (2005). Crop-to-weed introgression has impacted allelic composition of johnsongrass populations with and without recent exposure to cultivated sorghum. *Molecular Ecology* 14, 2143–2145. doi: 10.1111/j.1365-294X.2005.02579.x
- Murdock, G. P. (1959). *Africa: its peoples and their culture history* (New York: McGraw-Hill).
- Mutegi, E., Sagnard, F., Muraya, M., Kanyenji, B., Rono, B., Mwongera, C., et al. (2010). Ecogeographical distribution of wild, weedy and cultivated *Sorghum bicolor* (L.) Moench in Kenya: implications for conservation and crop-to-wild gene flow. *Genet. Resour. Crop Evol.* 57, 243–253. doi: 10.1007/s10722-009-9466-7
- Mutegi, E., Sagnard, F., Semagn, K., Deu, M., Muraya, M., Kanyenji, B., et al. (2011). Genetic structure and relationships within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. *Theor. Appl. Genet.* 122, 989–1004. doi: 10.1007/s00122-010-1504-5
- Mutegi, E., Sagnard, F., Labuschagne, M., Herselman, L., Semagn, K., Deu, M., et al. (2012). Local scale patterns of gene flow and genetic diversity in a crop-wild-weedy complex of sorghum (*Sorghum bicolor* (L.) Moench) under traditional agricultural field conditions in Kenya. *Conserv. Genet.* 13, 1059–1071. doi: 10.1007/s10592-012-0353-y
- Ng'uni, D., Geleta, M., Fatih, M., and Bryngelsson, T. (2010). Phylogenetic analysis of the genus *Sorghum* based on combined sequence data from cpDNA regions and ITS generate well-supported trees with two major lineages. *Ann. Bot.* 105, 471–480. doi: 10.1093/aob/mcp305
- Nwanze, K. F., Seetharama, N., Sharma, H. C., and Stenhouse, J. W. (1995). Biotechnology in pest management: Improving resistance in sorghum to insect pests. *Afr. Crop Sci. J.* 3, 209–215.
- Ohadi, S., Hodnett, G., Rooney, W., and Bagavathiannan, M. (2017). Gene Flow and its Consequences in Sorghum spp. *Crit. Rev. Plant Sci.* 36, 367–385. doi: 10.1080/07352689.2018.1446813
- Okeno, J. A., Mutegi, E., De Villiers, S., Wolt, J. D., and Misra, M. K. (2012). Morphological Variation in the Wild-Weedy Complex of *Sorghum bicolor* In Situ in Western Kenya: Preliminary Evidence of Crop-to-Wild Gene Flow? *Int. J. Plant Sci.* 173, 507–515. doi: 10.1086/665266
- Papa, R., Acosta, J., Delgado-Salinas, A., and Gepts, P. (2005). A genome-wide analysis of differentiation between wild and domesticated *Phaseolus vulgaris* from Mesoamerica. *Theor. Appl. Genet.* 111, 1147–1158. doi: 10.1007/s00122-005-0045-9
- Paterson, A. H., Schertz, K. F., Lin, Y. R., Liu, S. C., and Chang, Y. L. (1995). The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. *Proc. Natl. Acad. Sci. U. S. A.* 92, 6127–6131. doi: 10.1073/pnas.92.13.6127
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457, 551–556. doi: 10.1038/nature07723

- Piper, J. K., and Kulakow, P. A. (1994). Seed yield and biomass allocation in Sorghum bicolor and F1 and backcross generations of S. bicolor x S. halepense hybrids. *Can. J. Bot.* 72, 468–474. doi: 10.1139/b94-062
- Prescott-Allen, C., and Prescott-Allen, R. (1986). *The first resource: wild species in the North American economy* (New Haven: Yale University Press). doi: 10.2307/j.ctt211qvck
- Price, H. J., Dillon, S. L., Hodnett, G., Rooney, W. L., Ross, L., and Johnston, J. S. (2005a). Genome evolution in the genus *Sorghum* (Poaceae). *Ann. Bot.* 95, 219–227. doi: 10.1093/aob/mci015
- Price, H. J., Hodnett, G. L., Burson, B. L., Dillon, S. L., and Rooney, W. L. (2005b). A *Sorghum bicolor* x *S. macrospermum* hybrid recovered by embryo rescue and culture. *Aust. J. Bot.* 53, 579–582. doi: 10.1071/BT04213
- Price, H. J., Hodnett, G. L., Burson, B. L., Dillon, S. L., Stelly, D. M., and Rooney, W. L. (2006). Genotype Dependent Interspecific Hybridization of Sorghum bicolor. *Crop Sci.* 46, 2617–2622. doi: 10.2135/cropsci2005.09.0295
- Rich, P. J., Grenier, C., and Ejeta, G. (2004). Striga resistance in the wild relatives of sorghum. *Crop Sci.* 44, 2221–2229. doi: 10.2135/cropsci2004.2221
- Rosenow, D., and Clark, L. (1982). "Drought tolerance in sorghum [through breeding]," in *Proceedings of the annual corn and sorghum industry research conference-American Seed Trade Association* (Corn and Sorghum Research Conference (USA): Corn and Sorghum Division).
- Rowley-Conwy, P. A., Deakin, W. J., and Shaw, C. H. (1997). Ancient DNA from archaeological sorghum (*Sorghum bicolor*) from Qasr Ibrim, Nubia: Implications for domestication and evolution and a review of archaeological evidence. *Sahara* 9, 23–36. doi: 10.1007/978-1-4757-6730-8_5
- Sagnard, F., Deu, M., Dembélé, D., Leblois, R., Touré, L., Diakité, M., et al. (2011). Genetic diversity, structure, gene flow and evolutionary relationships within the Sorghum bicolor wild-weedy-crop complex in a western African region. *Theor. Appl. Genet.* 123, 1231. doi: 10.1007/s00122-011-1662-0
- Sangduen, N., and Hanna, W. (1984). Chromosome and fertility studies on reciprocal crosses between two species of autotetraploid Sorghum bicolor (L.) Moench and S. halepense (L.) Pers. *J. Hered.* 75, 293–296. doi: 10.1093/oxfordjournals.jhered.a109936
- Sasaki, T., and Antonio, B. A. (2009). Plant genomics: Sorghum in sequence. *Nature* 457:547. doi: 10.1038/457547a
- Schmidt, J. J., Pedersen, J. F., Bernards, M. L., and Lindquist, J. L. (2013). Rate of shattercane x sorghum hybridization in situ. *Crop Sci.* 53, 1677–1685. doi: 10.2135/cropsci2012.09.0536
- Schreiber, M., Himmelbach, A., Borner, A., and Mascher, M. (2019). Genetic diversity and relationship between domesticated rye and its wild relatives as revealed through genotyping-by-sequencing. *Evol. Appl.* 12 (1), 66–77. doi: 10.1111/eva.12624
- Sharma, H. C., and Franzmann, B. A. (2001). Host-plant preference and oviposition responses of the sorghum midge, *Stenodiplosis sorghicola* (Coquillett) (Dipt., Cecidomyiidae) towards wild relatives of sorghum. *J. Appl. Entomol.* 125, 109–114. doi: 10.1046/j.1439-0418.2001.00524.x
- Snowden, J. D. (1955). *The cultivated races of sorghum* (London: Adlard).
- Spangler, R., Zaitchik, B., Russo, E., and Kellogg, E. (1999). Andropogoneae evolution and generic limits in Sorghum (Poaceae) using ndhF sequences. *Syst. Bot.* 24, 267–281. doi: 10.2307/2419552
- Spangler, R. E. (2003). Taxonomy of Sarga, Sorghum and Vacoparis (Poaceae: Andropogoneae). *Aust. Syst. Bot.* 16, 279–299. doi: 10.1071/SB01006
- Stenhouse, J., Rao, K., Reddy, V., and Rao, S. (1997). "Sorghum", in *Biodiversity in trust: conservation and use of plant genetic resources in CGIAR centres*. Eds. D. Fuccillo, L. Sears, and P. Stapleton (Cambridge, United Kingdom: Cambridge University Press), 308–324.
- Sun, Y., Suksayretup, K., Kirkham, M. B., and Liang, G. H. (1991). Pollen tube growth in reciprocal interspecific pollinations of Sorghum bicolor, and S. versicolor. *Plant Breed.* 107, 197–202.
- Sun, Y., Skinner, D., Liang, G., and Hulbert, S. (1994). Phylogenetic analysis of Sorghum and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theor. Appl. Genet.* 89, 26–32. doi: 10.1007/BF00226978
- Tao, Y., Mace, E. S., Tai, S., Cruickshank, A., Campbell, B. C., Zhao, X., et al. (2017). Whole-genome analysis of candidate genes associated with seed size and weight in Sorghum bicolor reveals signatures of artificial selection and insights into parallel domestication in cereal crops. *Front. Plant Sci.* 8:1237:1237. doi: 10.3389/fpls.2017.01237
- Tesso, T., Kapran, I., Grenier, C., Snow, A., Sweeney, P., Pedersen, J., et al. (2008). The potential for crop-to-wild gene flow in sorghum in Ethiopia and Niger: a geographic survey. *Crop Sci.* 48, 1425–1431. doi: 10.2135/cropsci2007.08.0441
- Tsukiboshi, T., Koga, H., Uematsu, T., and Shimanuki, T. (1998). "Resistance of sorghum and sudangrass to ergot caused by Claviceps sp. and the cultural control of the disease", in *Bulletin of the National Grassland Research Institute*. Ed. Norinsho Sochi Shikenjo (Nishianasuno, Japan: National Grassland Research Institute), 28–35.
- Tyack, N., Dempewolf, H., Redden, R., Yadav, S., Maxted, N., Dulloo, M., et al. (2015). "The economics of crop wild relatives under climate change," in *Crop Wild Relatives and Climate Change* (Hoboken, New Jersey: John Wiley & Sons), 281–291.
- USDA and Agricultural Research Service, National Plant Germplasm System (2020). *Germplasm Resources Information Network (GRIN-Taxonomy)* (Beltsville, Maryland: National Germplasm Resources Laboratory). Available at: <https://npgsweb.ars-grin.gov/gringlobal/taxonomydetail.aspx?id=489172> (Accessed 25th March 2020).
- Venkateswaran, K., Muraya, M., Dwivedi, S. L., and Upadhyaya, H. D. (2014). "Wild sorghums-Their potential use in crop improvement", in *Genetics, genomics and breeding of sorghum*. Eds. Y. Wang, H. D. Upadhyaya and K. Chittaranjan (Florida, USA: CRC Press), 78–111.
- Venkateswaran, K., Elangovan, M., and Sivaraj, N. (2019a). "Origin, Domestication and Diffusion of Sorghum bicolor," in *Breeding Sorghum for Diverse End Uses*. Eds. C. Aruna, K.B.R.S. Visarada, B. V. Bhat and V. A. Tonapi (Cambridge, United Kingdom: Woodhead Publishing), 15–31.
- Venkateswaran, K., Sivaraj, N., Pandravada, S. R., Reddy, M. T., and Babu, B. S. (2019b). "Classification, Distribution and Biology", in *Breeding Sorghum for Diverse End Uses*. Eds. C. Aruna, K.B.R.S. Visarada, B. V. Bhat and V. A. Tonapi (Cambridge, United Kingdom: Woodhead Publishing), 33–60.
- Venkateswaran, K. (2003). *Diversity analysis and identification of sources resistant to downy mildew, shoot fly and stem borer in wild sorghums* Ph.D. thesis, (Hyderabad, India: Osmania University).
- Viaene, N. M., and Abawi, G. S. (1998). Management of Meloidogyne hapla on Lettuce in Organic Soil with Sudangrass as a Cover Crop. *Plant Dis.* 82, 945–952. doi: 10.1094/PDIS.1998.82.8.945
- Wang, Y. H., Upadhyaya, H. D., and Dweikat, I. (2015). "Sorghum," in *Genetic and Genomic Resources for Grain Cereals Improvement*. Eds. M. Singh and H. D. Upadhyaya (Cambridge, United Kingdom: Academic Press).
- Werle, R., Schmidt, J. J., Laborde, J., Tran, A., Creech, C. F., and Lindquist, J. L. (2014). Shattercane x ALS-tolerant sorghum F1 hybrid and shattercane interference in ALS-tolerant Sorghum. *J. Agric. Sci.* 6, 159–165.
- Wiersema, J. H., and Dahlberg, J. (2007). The nomenclature of Sorghum bicolor (L.) Moench (Gramineae). *Taxon* 56, 941–946. doi: 10.2307/25065876
- Winchell, F., Stevens, C. J., Murphy, C., Champion, L., and Fuller, D. Q. (2017). Evidence for Sorghum Domestication in Fourth Millennium BC Eastern Sudan Spikelet Morphology from Ceramic Impressions of the Butana Group. *Curr. Anthropol.* 58, 673–683. doi: 10.1086/693898
- Wooten, D. R. (2001). *The use of Sorghum propinquum to enhance agronomic traits in sorghum* Ph.D. thesis, [College Station (TX)]: (Texas A&M University).
- Yu, J., and Tuinstra, M. R. (2001). Genetic analysis of seedling growth under cold temperature stress in grain sorghum. *Crop Sci.* 41 (5), 1438–1443. doi: 10.2135/cropsci2001.4151438x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Ananda, Myrans, Norton, Gleadow, Furtado and Henry. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Characterization, Genetic Analyses, and Identification of QTLs Conferring Metabolic Resistance to a 4-Hydroxyphenylpyruvate Dioxygenase Inhibitor in Sorghum (*Sorghum bicolor*)

OPEN ACCESS

Edited by:

Mahalingam Govindaraj,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

Reviewed by:

Fernando Martinez,
University of Seville, Spain
Bernardo Ordas,
Consejo Superior de Investigaciones
Científicas (CSIC), Spain

*Correspondence:

Mithila Jugulam
mithila@ksu.edu

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 19 August 2020

Accepted: 09 November 2020

Published: 09 December 2020

Citation:

Pandian BA, Varanasi A,
Vennapusa AR, Sathishraj R, Lin G,
Zhao M, Tunnell M, Tesso T, Liu S,
Prasad PV and Jugulam M (2020)
Characterization, Genetic Analyses,
and Identification of QTLs Conferring
Metabolic Resistance to a
4-Hydroxyphenylpyruvate
Dioxygenase Inhibitor in Sorghum
(*Sorghum bicolor*).
Front. Plant Sci. 11:596581.
doi: 10.3389/fpls.2020.596581

Balaji Aravindhan Pandian¹, Aruna Varanasi², Amaranatha R. Vennapusa¹,
Rajendran Sathishraj¹, Guifang Lin³, Mingxia Zhao³, Madison Tunnell¹, Tesfaye Tesso¹,
Sanzhen Liu³, P. V. Vara Prasad^{1,4} and Mithila Jugulam^{1*}

¹ Department of Agronomy, Kansas State University, Manhattan, KS, United States, ² Bayer Crop Science, St. Louis, MO, United States, ³ Department of Plant Pathology, Kansas State University, Manhattan, KS, United States, ⁴ Sustainable Intensification Innovation Lab, Kansas State University, Manhattan, KS, United States

Postemergence grass weed control continues to be a major challenge in grain sorghum [*Sorghum bicolor* (L.) Moench], primarily due to lack of herbicide options registered for use in this crop. The development of herbicide-resistant sorghum technology to facilitate broad-spectrum postemergence weed control can be an economical and viable solution. The 4-hydroxyphenylpyruvate dioxygenase-inhibitor herbicides (e.g., mesotrione or tembotrione) can control a broad spectrum of weeds including grasses, which, however, are not registered for postemergence application in sorghum due to crop injury. In this study, we identified two tembotrione-resistant sorghum genotypes (G-200, G-350) and one susceptible genotype (S-1) by screening 317 sorghum lines from a sorghum association panel (SAP). These tembotrione-resistant and tembotrione-susceptible genotypes were evaluated in a tembotrione dose-response [0, 5.75, 11.5, 23, 46, 92 (label recommended dose), 184, 368, and 736 g ai ha⁻¹] assay. Compared with S-1, the genotypes G-200 and G-350 exhibited 10- and seven fold more resistance to tembotrione, respectively. To understand the inheritance of tembotrione-resistant trait, crosses were performed using S-1 and G-200 or G-350 to generate F₁ and F₂ progeny. The F₁ and F₂ progeny were assessed for their response to tembotrione treatment. Genetic analyses of the F₁ and F₂ progeny demonstrated that the tembotrione resistance in G-200 and G-350 is a partially dominant polygenic trait. Furthermore, cytochrome P450 (CYP)-inhibitor assay using malathion and piperonyl butoxide suggested possible CYP-mediated metabolism of tembotrione in G-200 and G-350. Genotype-by-sequencing based quantitative trait loci (QTL) mapping revealed QTLs associated with tembotrione resistance in G-200 and G-350 genotypes. Overall,

the genotypes G-200 and G-350 confer a high level of metabolic resistance to tembotrione and controlled by a polygenic trait. There is an enormous potential to introgress the tembotrione resistance into breeding lines to develop agronomically desirable sorghum hybrids.

Keywords: *Sorghum bicolor*, resistance, QTL mapping, single gene inheritance, tembotrione

INTRODUCTION

Grain sorghum [*Sorghum bicolor* (L.) Moench ssp. *bicolor*] is one of the most versatile crops with multiple uses, including for food, feed, and fuel (Ciampitti and Prasad, 2019). Sorghum performs better than corn (*Zea mays*) under rainfed and low input conditions (Valadabad et al., 2000; Staggenborg et al., 2008). The US is the largest producer of grain sorghum in the world, and almost half of the US grain sorghum is produced in Kansas (USDA-NASS, 2020). Sorghum is primarily grown for cattle feed and ethanol production in the US, whereas it is a staple food for millions of people in Africa, India, and South America (Taylor et al., 2006; Dahlberg et al., 2012). Weed infestation, specifically grass weed species, pose a major problem in sorghum production and can reduce the crop yields up to 60% if left uncontrolled (Thompson et al., 2019; Dille et al., 2020). Palmer amaranth (*Amaranthus palmeri*), common waterhemp (*Amaranthus tuberculatus*), kochia (*Bassia scoparia*), common ragweed (*Ambrosia artemisiifolia*), and common lambsquarters (*Chenopodium album*) are the major broadleaf weeds and johnsongrass (*Sorghum halepense*), shattercane (*Sorghum bicolor* ssp. *verticilliflorum*), and large crabgrass (*Digitaria sanguinalis*) are major grass weeds found in grain sorghum fields (Stahlman et al., 2000; Smith et al., 2010). A wide range of postemergence (POST) herbicides are available to control broad-leaved weeds in sorghum. However, herbicide options for POST control of grasses are limited due to the susceptibility of sorghum to commonly used grass control herbicides (Thompson et al., 2019).

The 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitors (e.g., mesotrione or tembotrione) are widely used to control a broad spectrum of weeds including grasses in corn because it can effectively metabolize HPPD inhibitors (Williams and Pataky, 2010). However, these herbicides are not registered as POST in sorghum due to crop injury. Although these herbicides are widely used, to date only two weed species, i.e., Palmer amaranth and common waterhemp, have been documented to have evolved resistance to HPPD inhibitors (Heap, 2020). These herbicides inhibit the HPPD enzyme, which is important for the conversion of 4-hydroxyphenyl pyruvate to homogentisate, an intermediate in plastoquinone and tocopherol biosynthesis pathway in plants (Lee et al., 1998). Plastoquinone is essential for the carotenoid biosynthesis, which protects the chlorophyll by absorbing excited electrons released during photosynthesis. Depletion of carotenoids causes damage to the chlorophyll by photo-oxidation resulting in bleaching followed by necrosis and plant death (Dankov et al., 2009). HPPD inhibitors include four chemical families isoxazole, pyrazole, pyrazolone, and triketones, and were introduced in the 1980s for weed control (van Almsick, 2009).

Herbicide resistance in plants can be conferred by two major mechanisms: (1) target-site resistance (TSR): mutation(s) in the herbicide target gene leading to the reduced affinity of the target enzyme for herbicide binding or due to increased expression of target enzyme; (2) non-target site resistance (NTSR): increased metabolism or reduced absorption/translocation of herbicides (Gaines et al., 2020). Metabolism of HPPD inhibitors by cytochrome P450 enzyme (CYPs) activity is the most common mechanism of resistance found in crops as well as weeds (Ahrens et al., 2013). Nonetheless, increased expression of *HPPD* gene has also been reported in some biotypes of Palmer amaranth (Nakka et al., 2017). Recently, a modified *HPPD* gene from *Pseudomonas fluorescens* and *Avena sativa* which is insensitive to HPPD inhibitors was used to develop transgenic soybeans (*Glycine max*) resistant to HPPD inhibitors by Bayer Crop Science (Matringe et al., 2005; Dreesen et al., 2018) and Syngenta (Hawkes et al., 2016), respectively. Dupont-Pioneer used an insensitive shuffled variant of corn *HPPD* gene that confers a high level of resistance to HPPD inhibitors in soybean (Siehl et al., 2014).

CYPs are one of the largest enzyme families involved in xenobiotic metabolism in microorganisms, insects, plants, and humans imparting resistance, respectively, to antibiotics, insecticide, herbicide, and drugs (Pandian et al., 2020). The activity of CYPs can be inhibited using several chemical compounds: 1-aminobenzo-triazole (ABT), tetracyclis (TET), piperonyl butoxide (PBO), tridiphane, and organophosphate insecticides such as malathion and phorate (Siminszky, 2006; Busi et al., 2017). Treatment with CYP inhibitors before herbicide application will competitively reduce the CYP activity resulting in decreased metabolism of herbicide, thereby reducing the level of resistance (Siminszky, 2006). CYP inhibitors have been widely used to determine metabolic resistance to herbicides in several plant species.

Specifically, malathion and PBO were used to demonstrate the inhibition of CYP activity and the reversal of crop tolerance to HPPD inhibitors in corn (Ma et al., 2013; Oliveira et al., 2018).

Development of sorghum hybrids resistant to HPPD inhibitors will provide POST herbicide options to control grass weeds (Thompson et al., 2019). Tembotrione is a triketone herbicide which has broad-spectrum activity including grass weeds. Furthermore, the efficacy of tembotrione is high on grass weeds compared with other triketones (Ahrens et al., 2013). Mesotrione, a triketone herbicide similar to tembotrione, is registered for pre-emergence (PRE) use in sorghum but not as POST; however, tembotrione is not registered for PRE or POST usage in sorghum. We have used sorghum association panel (SAP) composed of homozygous sorghum genotypes representing all cultivated races from diverse geographic regions including widely used US breeding lines. We hypothesize

that screening diverse genotypes from the SAP will facilitate the identification of genotypes resistant to tembotrione; such resistance, similar to maize, is associated with CYP-mediated metabolism. The specific objectives of this research were to identify and characterize sorghum genotypes with resistance to tembotrione, to investigate the inheritance and mechanism of resistance to tembotrione, and to identify genetic loci conferring tembotrione resistance.

MATERIALS AND METHODS

Plant Materials

Sorghum genotypes from the SAP (Casa et al., 2008) were used in this study. A commercial sorghum hybrid Pioneer 84G62 and a corn inbred B73 (naturally resistant to tembotrione) were also used for comparison.

In vitro Screening

Sorghum genotypes (~317) from SAP along with Pioneer 84G62 and B73 were used for initial screening with tembotrione under *in vitro* conditions. Seeds of all genotypes were germinated in plastic Petri dishes (100 mm diameter × 20 mm height) containing 0.8% w/v solidified agar medium (PhytoTech Laboratories, Lenexa, KS, United States). Seeds were surface sterilized with 2% ethanol for 2 min followed by 5% (v/v) sodium hypochlorite for 15 min. Subsequently, seeds were rinsed two to three times with sterile distilled water before placing them on the agar medium. About 8–10 seeds were placed in each Petri dish for germination and incubated in a growth chamber maintained at 24°C with 16/8 h (day/night) photoperiod under a photosynthetic flux of 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ (daylight fluorescent tubes). On germination, seedlings at three-leaf stage were transferred to culture vessels (PhytoTech Laboratories) containing solidified agar supplemented with 0.25 μM molecular grade tembotrione (Sigma-Aldrich, St. Louis, MO, United States). All transplanted culture vessels were incubated in the same growth chamber, maintained at the same conditions as indicated previously. The experiment was conducted with four to eight biological replicates (two culture vessels with two to four plants in each culture vessel). The response of genotypes to tembotrione treatment was evaluated visually (percent injury) at 2 and 4 weeks after treatment (WAT) based on a 0–100% rating scale (0% is no injury and 100% is complete death) (Abit et al., 2009).

Whole Plant Assay

Ten sorghum genotypes (Supplementary Table S1) that exhibited minimum injury and S-1 that was found highly susceptible to tembotrione under *in vitro* conditions were tested along with Pioneer 84G62 for their response to tembotrione under greenhouse conditions. The seeds of sorghum genotypes were planted in square pots (15 × 15 × 15 cm) filled with a potting mixture (ProMix Ultimate; Premier Tech Horticulture, Mississauga, Ontario, Canada). The seedlings at three-leaf stage (Roozeboom and Prasad, 2019) were transplanted in square pots (6 × 6 × 6 cm) and grown in a greenhouse maintained at 25/20°C, 15/9 h day/night photoperiod with a

photosynthetic photon flux density of 750 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and relative humidity of $60 \pm 10\%$. The plants were fertilized (Miracle GRO® all-purpose plant food; ScottsMiracle-Gro, Marysville, OH, United States) as needed. The sorghum seedlings at five-leaf stage (Roozeboom and Prasad, 2019) were treated with tembotrione (Laudis; Bayer Crop Science, St. Louis, MO, United States) at 92 g ai ha⁻¹ (field recommended dose) with 0.25% methylated soy oil (Destiny; WinField United) using a bench-top track spray chamber (Generation III; De Vries Manufacturing, Hollandale, MN, United States) equipped with a single flat-fan nozzle (80015LP TeeJet tip; Spraying Systems, Wheaton, IL, United States) delivering 187 L ha⁻¹. Each plant was considered as an experimental unit, eight replications were used for each genotype. The response of sorghum genotypes to tembotrione treatment was evaluated by visual injury rating as described previously (Abit et al., 2009). The above-ground plant biomass was harvested 3 WAT and dried in an oven at 60°C for 72 h. The weight of dried biomass was recorded as described later in a separate section. The experiment was repeated two times following the same procedure and growth conditions.

Dose-Response Assay

Two genotypes, i.e., G-200 and G-350, that exhibited the least injury, and one highly susceptible genotype S-1 that exhibited the highest injury to tembotrione treatment identified from *in vitro* and whole plant assays, along with Pioneer 84G62 were tested in a tembotrione dose-response study to determine the level of resistance. The sorghum genotypes were treated with tembotrione at 0, 5.75, 11.5, 23, 46, 92, 184, 368, and 736 g ai ha⁻¹. The experiment was conducted following the same plant growth conditions and herbicide application procedure as described previously in the whole plant assay. The experiment was conducted in a completely randomized design with four replications and repeated twice. The above-ground plant biomass reduction was measured as described previously.

Field Testing

On confirmation of the level of resistance to tembotrione in the greenhouse, the tembotrione-resistant sorghum genotypes G-200 and G-350 were evaluated in comparison with S-1 (susceptible) and Pioneer 84G62 (commercial hybrid) under field conditions. Experiments were conducted in summer of 2017 at two KSU research sites: Ashland Bottoms Research Farm, Manhattan (Reading silt loam soil type; Pachic Agriustolls taxonomic class); and Agricultural Research Center, Hays (Harney silt loam soil type; Typic Agriustolls taxonomic class). S-Metolachlor at 2 kg ai ha⁻¹ was applied as a pre-emergence herbicide to all plots at both sites to suppress existing weeds in the field before planting sorghum. Seeds of sorghum genotypes were planted in both locations on June 6, 2017 with a 76-cm space between rows and 7.6-cm space between plants, and 2.5 cm planting depth at a rate of 172,000 seeds ha⁻¹ (Abit et al., 2011). The experimental plots were 3 m wide and 6 m long with four rows; the resistant or susceptible genotypes were planted in the middle two rows along with two border rows planted with Pioneer 84G62 to avoid herbicide drift from nearby treatments.

POST application of tembotrione was made to individual plots when the sorghum plants reached five-leaf stage (Roozeboom and Prasad, 2019). Tembotrione treatments included 0, 92, 184, and 368 g ai ha⁻¹. Herbicides were applied using a CO₂-powered backpack-type research sprayer equipped with TurboTee 11002 nozzles calibrated to deliver 140 L ha⁻¹ at 234 kPa. Experiments were conducted in a randomized complete block design with factorial arrangement with sorghum genotype and herbicide dose as the two factors. All treatments were replicated four times at each site. Sorghum response to herbicide treatments was visually assessed 1, 2, 4, and 8 WAT using a scale of 0 (no visible injury) to 100% (plant death) compared with the non-treated plants. At physiological maturity, grain weight was measured for each genotype. Three sorghum heads from each replication were randomly collected for each genotype separately from all treatments and dried in an oven at 60°C for 1 week. The dried sorghum heads from each plant were subsequently threshed to determine grain yield from a single plant.

Generation and Evaluation of F₁ and F₂ Progeny

To study the inheritance and mapping of tembotrione resistance, direct and reciprocal crosses were performed using tembotrione-resistant (G-200 and G-350) and tembotrione-susceptible (S-1) genotypes in a crossing nursery at KSU research farm, Ashland Bottoms, KS. The crosses were made using the plastic bag method (Rakshit and Bellundagi, 2019). The F₁ seeds were harvested from individual plants. The F₁ progeny from S-1 × G-200 and S-1 × G-350 were evaluated in a tembotrione dose-response assay by treating the plants with 0, 23, 46, 92, 184, and 368 g ai ha⁻¹ of tembotrione. A total of 10–12 F₁ plants from each cross (S-1 × 200 and S-1 × 350) per dose were treated and the true F₁ plants were differentiated from the selfed plants by their response to tembotrione; hence, the susceptible (S-1) was used as female parent, and the plants that were derived by selfing would be killed at field recommended dose or higher. In addition, the selfed plants that survived at low doses were identified by parental phenotype and vigor and discarded.

Each plant was considered as an experimental unit with eight replications per dose. The same procedure, as described previously, was followed for tembotrione dose-response assay of F₁ progeny. Three F₁ plants per cross that exhibited resistance to tembotrione were selected to generate F₂ seeds by self-pollination.

The F₂ progeny were evaluated under greenhouse conditions with a single dose of tembotrione to determine the segregation of resistant and susceptible plants. Approximately 150 seedlings from a single F₂ family (total of two F₂ families) along with the parents were raised in the greenhouse (as described previously under the same growth conditions). The seedlings (five-leaf stage) were treated with 276 g ai ha⁻¹ of tembotrione following the same procedure as described previously. The response of F₂ plants was assessed by visual injury rating (as described previously) at 2 and 3 WAT (Abit et al., 2009). Further, plants were grouped as highly injured/dead (susceptible) or minor/no symptoms (resistant) at 4 WAT in comparison with the parental genotypes. In addition,

total leaf chlorophyll index was estimated in parents and F₂ progeny on 3 and 4 WAT. Chlorophyll index was measured at three different spots on the leaf blade along the length of the youngest fully opened leaf using a self-calibrating soil plant analysis development (SPAD) chlorophyll meter (Konica Minolta SPAD 502 Chlorophyll Meter, Chiyoda City, Tokyo, Japan). The chlorophyll index obtained from the three spots were averaged and considered as a total leaf chlorophyll index. However, the leaf chlorophyll index was recorded from the second run of S-1 × G-200 F₂ evaluation which was used for the quantitative trait loci (QTL) mapping experiment (described later in a separate section).

HPPD-Gene Sequencing

The *HPPD* gene from G-200, G-350, and S-1 were sequenced to determine if any target site alterations confer resistance to tembotrione. Leaf tissue (three-leaf stage plants) was collected from three plants (biological replicates) of each genotype grown in the greenhouse as described previously and under similar growth conditions. The genomic DNA was extracted using GeneJET Plant Genomic DNA Purification Mini Kit (Thermo Fisher Scientific, Waltham, MA, United States) following the manufacturer's instructions. The concentration of the DNA samples was quantified using NanoDrop (Thermo Fisher Scientific). The sorghum *HPPD* gene ~2 kb was amplified using the primers Sg_HPPD F (5' GACACGATGAATGCCCATGC 3') and Sg_HPPD R (5' AGAGAGATGACAGTACAGTGTGTGT 3') designed from Sobic.002G104200.1 in the sorghum reference genome V3.1.1 (McCormick et al., 2017). PCR was performed using T100 Thermal Cycler (Bio-Rad, Hercules, CA, United States). The PCR mixture contained 50–80 ng of gDNA, 0.5 μM each of forward, reverse primer, and 1 × of GoTaq G2 Green Master Mix (Promega, Madison, WI, United States). PCR amplification was done using the following PCR cycling conditions: initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 45 s, extension at 72°C for 45 s, and final extension at 72°C for 7 min. The PCR products were analyzed in 1.5% agarose gel to confirm the targeted amplicon size and purified using GeneJET PCR Purification Kit (Thermo Fisher Scientific, Waltham, MA, United States). The PCR purified samples were sequenced by Sanger sequencing service provided by GENEWIZ (South Plainfield, New Jersey, United States). The sequences were aligned using Clustal Omega multiple sequence alignment tool (EMBL-EBI) to check for the mutations.

CYP-Inhibitor Study

To determine if CYP-mediated metabolism of tembotrione confers resistance in G-200 and G-350 genotypes, experiments were conducted using two CYP inhibitors, malathion and PBO. The sorghum genotypes G-200, G-350, and S-1, along with Pioneer 84G62 and a corn genotype B73, were grown in the greenhouse (as described previously and under similar growth conditions). Malathion (Spectracide malathion insect spray concentrate; Spectrum Brands) at 0, 2,000, and 4,000 g ai ha⁻¹ or PBO (Thermo Fisher Scientific, Waltham, MA, United States) at 4,500 g ai ha⁻¹ along with 0.25% non-ionic

surfactant (NIS) was applied 1 h before tembotrione treatment. Soil drenching of 5 mM malathion 24 h after primary application as a booster dose was given only for the malathion treatments. Tembotrione was applied at 0, 92, 184, and 368 g ai ha⁻¹ with 0.25% methylated soy oil. All the treatments were arranged in a factorial design. The same procedure, as mentioned in the earlier tembotrione dose-response assay, was followed for chemical treatments (malathion, PBO, and tembotrione) and data collection.

Genotyping by Sequencing (GBS)

A total of 150 plants of a single F₂ family derived from S-1 × G-200 (mentioned as Run-2 in **Table 4**), along with parents, were grown in the greenhouse as described previously and under the same growth conditions. An equal amount (two 2-cm leaf bits; ~150 mg) of leaf tissue was collected from all plants in 96-deep-well plates. One 3.2-mm stainless steel bead was added to each well and the leaf tissue was ground for 3 min at 20 cycles per second to obtain fine powder in a Mixer Mill (Retsch GmbH, Haan, North Rhine-Westphalia, Germany). Genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (Bai et al., 1999) with minor modifications. The DNA concentration in the extracted samples was quantified by FLUOstar Omega microplate reader (BMG LABTECH, Ortenberg, Baden-Württemberg, Germany) using a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, Grand Island, NY, United States). Each sample was normalized to contain 10 ng/μl DNA using QIAgility Liquid Handling System (Qiagen, Germantown, MD, United States) for library construction. Approximately 150 ng of genomic DNA of each sample was used to construct a library following the tGBS protocol (Ott et al., 2017) with modifications, and the DNA library was sequenced on the HiSeqX 10 platform at Novogene Corporation, Sacramento, CA, United States. Sequencing reads were trimmed and de-barcoded using the pipeline described in the previous tGBS study (Ott et al., 2017). Clean reads of each sample were aligned to *Sorghum bicolor* genome (GenBank accession GCA_000003195.3) (Paterson et al., 2009) using BWA v0.7.12-r1039 (Li and Durbin, 2009) and unique mapped reads were retained for variant discovery using HaplotypeCaller in GATK v4.1 (McKenna et al., 2010). GATK SelectVariants with parameters (-select-type SNP -restrict-alleles-to BIALLELIC -select "QD ≥ 10.0" -select "DP ≥ 200.0") was applied to filter variants. The SNPs were converted to ABH format ("A" represents resistant parent allele, "B" represents susceptible allele, and "H" represents heterozygous allele) and only polymorphic SNPs between the R and S genotypes were retained using a custom-made Microsoft Excel template. The filtered polymorphic SNPs were used for the construction of a linkage map and QTL analysis.

Linkage and QTL Mapping

The linkage map was obtained using QTL IciMapping (version 4.5). The grouping and ordering of 606 polymorphic SNP markers were carried out using the regression mapping algorithm RECORD (REcombination Counting and ORDERing) based on recombination events between adjacent markers. Further,

ripping was done for fine-tuning of the ordered markers on their respective chromosomes by the sum of adjacent recombination fractions (SARF) algorithm with a default window size. The QTL mapping for recovery (RE) and visual injury (VI) was performed using the inclusive composite interval mapping (ICIM) method with additive assumption was performed using the QTL IciMapping (version 4.5) (Meng et al., 2015). The logarithm of the odds (LOD) significance thresholds ($P < 0.05$) were determined by running 1,000 permutations (Churchill and Doerge, 1994). Previously reported QTLs for similar regions were obtained from sorghum QTL Atlas (Mace et al., 2019). The QTLs (q) were named based on the trait abbreviation followed by the chromosome number.

Statistical Analysis

Dry biomass (% of non-treated) was calculated following the formula:

$$\text{Dry biomass (\% of non - treated)} = \frac{\text{Biomass of individual plant (g)}}{\text{Average biomass of the non - treated plants of the genotype (g)}} \times 100$$

Tembotrione dose-response data expressed as dry biomass (% of non-treated) or percent injury were subjected to non-linear regression analysis using a three- or four-parameter log-logistic model using a "drc" (Ritz and Streibig, 2005) package in R (Development R Core Team, 2013) following Knezevic et al. (2007) and Shyam et al. (2019) to estimate GR₅₀ (dose required for 50% growth reduction) or ID₅₀ (dose required for 50% visual injury). A "Lack-of-fit" test was performed using the "model fit" function of "drc" to assess the fit of data to various regression models. Differences between the estimated GR₅₀ or ID₅₀ values were tested with each other by *t*-test using the "compParm" function in the "drc" package. The dose-response curves were generated using the "plot" function in the "drc" package.

ANOVA was performed following Fisher's LSD test to separate means and significance at $P \leq 0.05$ using the "agricole" package in R (de Mendiburu, 2014). The plots were generated using the "R" package "ggplot2" (Wickham and Wickham, 2007). A χ^2 goodness-of-fit test (Cochran, 1952) was used to fit to a single dominant gene by comparing the observed and expected segregation frequencies of tembotrione-resistant or tembotrione-susceptible plants.

RESULTS

Identification of Tembotrione-Resistant Genotypes

Among 317 genotypes from the SAP initially screened under *in vitro* (tissue culture) conditions, 10 genotypes showed $\leq 70\%$ tembotrione injury at 2 WAT with significant recovery by 4 WAT (**Supplementary Table S1**). One genotype, S-1, was

found highly susceptible to tembotrione as these plants did not survive tembotrione treatment. In response to the tembotrione application at 92 g ai ha⁻¹ (field recommended dose of tembotrione), under greenhouse conditions, out of the above 10 genotypes, only two, i.e., G-200 and G-350, showed the least injury at 2 WAT (73 and 70%, respectively) and the smallest dry biomass reduction at 3 WAT (25.36 and 25.77%, respectively) (Supplementary Table S2). In the treated plants, the tembotrione injury symptoms appeared by 2 WAT, after which in G-200 and G-350, the symptoms dissipated, and plants started to recover. However, leaf chlorosis and bleaching symptoms, typical of tembotrione injury, were visible on all 10 genotypes at a variable degree. The susceptible genotype, S-1, and Pioneer 84G62 (a commercially used sorghum hybrid) exhibited 100 and 90% injury with a biomass reduction of 2.69 and 11.63%, respectively (Supplementary Table S2).

Tembotrione Dose-Response Assay

Based on the GR₅₀ values the resistant genotypes, G-200 and G-350 required 215 and 154 g ai ha⁻¹ of tembotrione, respectively, for 50% growth reduction which is higher than the field recommended dose of 92 g ai ha⁻¹, whereas Pioneer 84G62 and S-1 required only 22 and 36 g ai ha⁻¹ for the same level of growth reduction. The GR₅₀ values of G-200, G-350, and Pioneer 84G62 were significantly different from S-1. Based on GR₅₀ values, compared with S-1, the genotypes G-200, G-350, and Pioneer 84G62 were ~10 ×, 7 ×, and 1.5 × more resistant, respectively, to tembotrione (Figure 1 and Table 1).

Field Testing

The response of sorghum genotypes, G-200 and G-350, along with S-1 and Pioneer 84G62, to POST application of tembotrione

TABLE 1 | Regression parameters describing the response of sorghum genotypes to tembotrione under greenhouse and field conditions.

Genotype	Greenhouse		Field	
	GR ₅₀ (SE)	RI (R/S)	ID ₅₀ (SE)	RI (R/S)
S-1	22.1 (3.4)	1	68.4 (14.3)	1.0
Pioneer 84G62	36.7 (5.6)	2	62.4 (51.0)	1.0
G-350	154.4 (68.8)**	7	96.6 (17.8)*	1.5
G-200	215.2 (160)*	10	129.6 (71.8)**	2.0

GR₅₀, dose required for 50% growth reduction; ID₅₀, dose required for 50% visual injury; SE, standard error; RI, resistance index; R/S, resistant/susceptible. Significantly different from S-1 at * $p \leq 0.05$, ** $p \leq 0.01$.

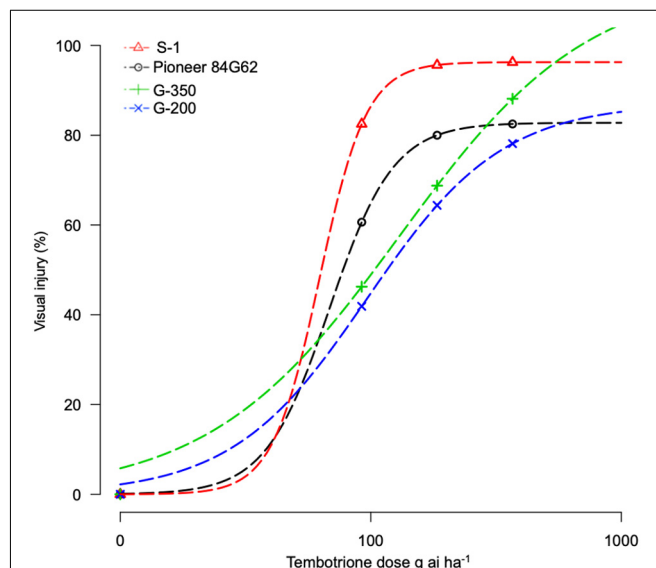


FIGURE 2 | Tembotrione dose-response curves representing the percent injury of S-1 (susceptible), Pioneer 84G62 (commercial hybrid), G-200, and G-350 (resistant genotypes) in response to different doses of tembotrione at 2 weeks after treatment under field conditions.

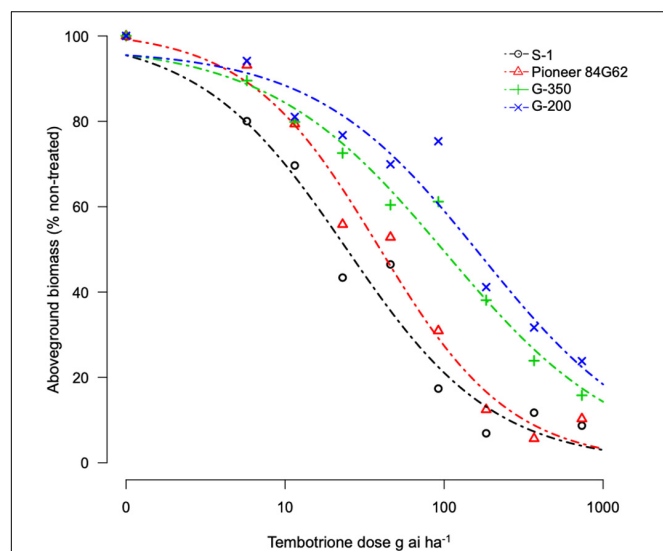


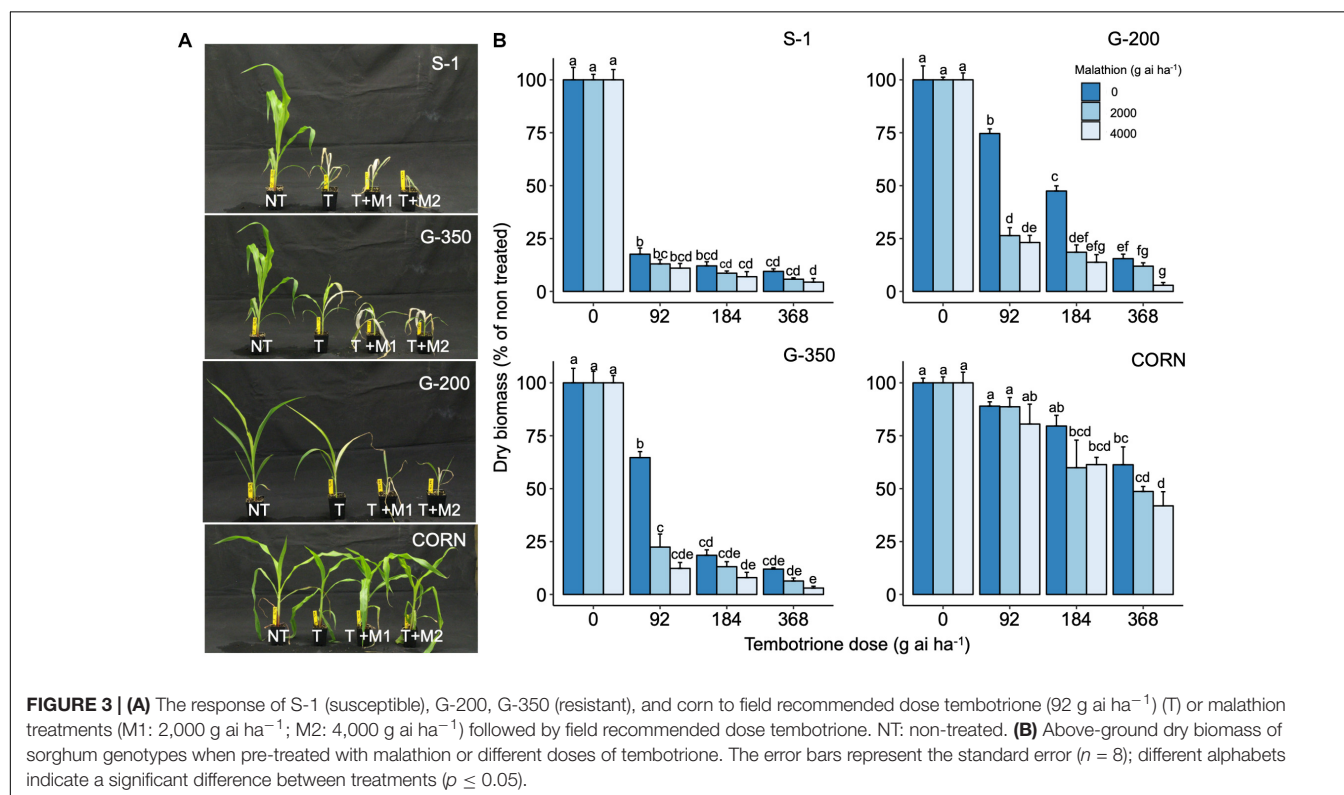
FIGURE 1 | Tembotrione dose-response curves obtained by non-linear regression analysis of above-ground dry biomass of S-1 (susceptible), Pioneer 84G62 (commercial hybrid), G-200, and G-350 (resistant) using the three-parameter log-logistic model.

was tested under field conditions at two sites, Manhattan and Hays, KS. Because site by herbicide dose interaction was non-significant, data for tembotrione injury and yield were pooled and averaged across the two sites. In response to the tembotrione (POST) application, sorghum genotypes showed leaf chlorosis and bleaching symptoms followed by necrotic lesions in susceptible plants. The tembotrione injury was visible on all four sorghum genotypes at 1 WAT. A dose-response curve was generated using the injury ratings at 2 WAT. Based on percent injury at 2 WAT and ID₅₀ values, the genotype G-200 required the highest dose (129 g ai ha⁻¹) of tembotrione followed by G-350 (96 g ai ha⁻¹) genotype. Susceptible genotypes S-1 (62 g ai ha⁻¹) and Pioneer 84G62 required the lowest dose of tembotrione (68 g ai ha⁻¹) that caused 50% injury (Figure 2 and Table 1). Days required for 50% recovery (ED₅₀) from herbicide damage was also calculated. The genotypes, G-200 and G-350, recovered with no injury symptoms by 6 WAT at all doses of tembotrione, whereas Pioneer 84G62 and S-1 took 8 weeks for recovery from 1 × and 2 × of application dose, and S-1 did

TABLE 2 | Recovery (ED_{50}) and single plant yield (% of non-treated) in response to different doses of tembotrione under field conditions (ED_{50} : days required for 50% recovery).

Herbicide dose (g ai ha ⁻¹)	ED_{50} (SE)				Yield (%)			
	S-1	Pioneer 84G62	G-350	G-200	S-1	Pioneer 84G62	G-200	G-350
0	—	—	—	—	100a	100a	100a	100a
92	28 (6)	23 (2)	22 (3)*	18 (1)*	53.7de	68.5cd	95.3ab	99.8a
184	30 (10)	25 (3)	21 (2)*	20 (1)*	19.4g	45.4cd	72.2cd	78.0bc
368	—	27 (1)	21 (3)	19 (1)	—	14.6g	25.3fg	33.3fg

Significantly different from S-1 at * $p \leq 0.05$. Numbers with different letters are significantly different; numbers with the same letters are not significantly different.



not survive $4 \times$ dose of tembotrione (Table 2). In response to $1 \times$ dose of tembotrione, no significant reduction in grain yield of G-200 and G-350 was found, whereas the susceptible genotypes, S-1 and Pioneer 84G62, showed ~ 40 and 30% grain yield reduction (Table 2) compared with no treatment. However, when treated with $2 \times$ and $4 \times$ doses of tembotrione, grain yields were significantly reduced in all genotypes (Table 2).

Mechanism of Tembotrione Resistance

Upon sequencing the *HPPD* gene from G-200, G-350, and S-1, no mutations were identified in the coding region of the *HPPD* gene (Supplementary Data 1), suggesting that no target site alterations confer resistance to tembotrione in G-200 or G-350. In response to malathion or PBO followed by tembotrione treatments, G-200 and G-350 exhibited significant biomass reduction compared with plants treated with tembotrione alone. Malathion or PBO without tembotrione treatment had no effect on the sorghum genotypes tested (Figures 3, 4). The corn

inbred line B73 (known to be resistant to tembotrione) did not show any significant biomass reduction at 92 or 184 g ai ha⁻¹ of tembotrione application (Figure 3A); however, it exhibited a significant reduction in biomass in response to pre-treatment with malathion followed by 184 and 368 g ai ha⁻¹ of tembotrione (Figure 3B). The genotype G-200 treated with malathion followed by 92, 184, and 368 g ai ha⁻¹ showed more than 50% reduction in biomass compared with plants treated only with tembotrione. There is no significant difference in biomass accumulation when treated with 2,000 or 4,000 g ai ha⁻¹ of malathion, except in malathion followed by 368 g ai ha⁻¹ tembotrione treatment (Figure 3B), whereas G-350 showed significant biomass reduction only when treated with malathion followed by 92 g ai ha⁻¹ tembotrione. The susceptible genotype, S-1, showed significant growth reduction at all doses of tembotrione or when pre-treated with malathion (Figure 3B). The corn B73 exhibited significant biomass reduction in response to PBO, followed by 92, 184, and 368 g

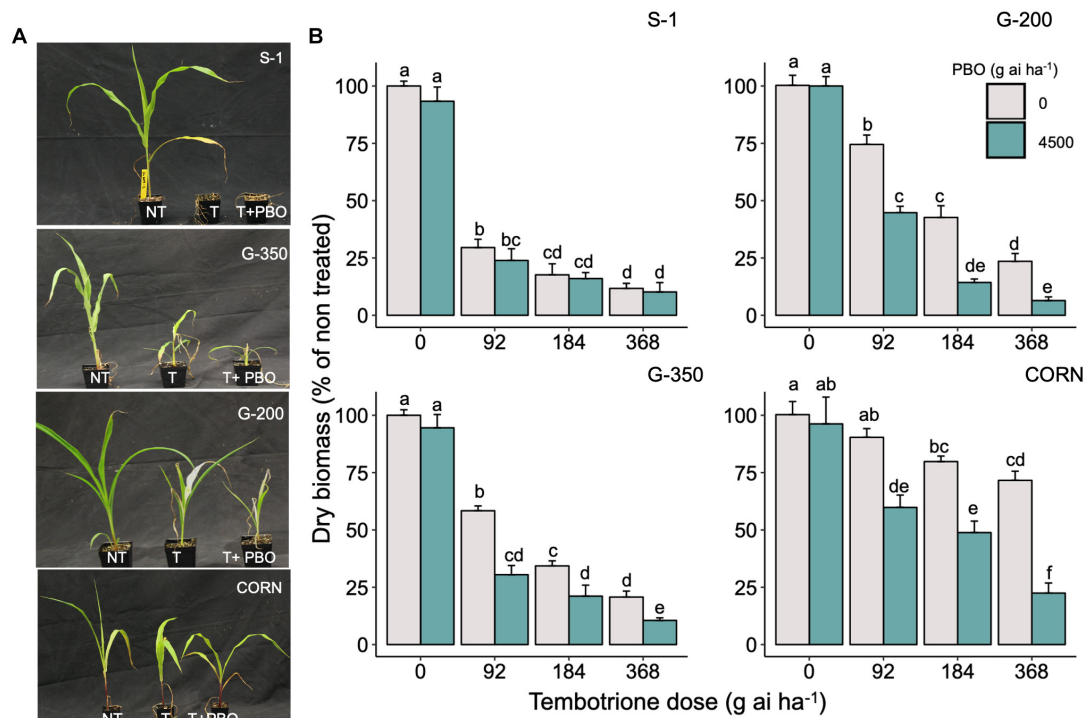


FIGURE 4 | (A) The response of S-1 (susceptible), G-200, G-350 (resistant), and corn to field recommended dose of tembotrione (92 g ai ha⁻¹) (T) or piperonyl butoxide (PBO) treatment (M1: 4,500 g ai ha⁻¹) followed by field recommended dose tembotrione. NT: non-treated. **(B)** Above-ground dry biomass of sorghum genotypes when pre-treated with PBO or different doses of tembotrione. The error bars represent the standard error ($n = 8$); different alphabets indicate a significant difference between treatments ($p \leq 0.05$).

ai ha⁻¹ of tembotrione treatment (Figures 4A,B). Both G-200 and G-350 genotypes also showed a significant reduction in biomass when pre-treated with PBO followed by all doses of tembotrione. The S-1 was susceptible to all treatments applied (Figure 4B).

Inheritance of Tembotrione Resistance

The F₁ progeny of S-1 × G-200 and S-1 × G-350 showed an intermediate response relative to parents when treated with several doses of tembotrione. The GR₅₀ of S-1 × G-200 and S-1 × G-350 were estimated at 104 and 117 g ai ha⁻¹, respectively, which were less than their respective tembotrione-resistant parents, i.e., G-200 (218 g ai ha⁻¹) and G-350 (172 g ai ha⁻¹) (Figures 5, 6 and Table 3), suggesting that tembotrione resistance is a partially dominant trait. The F₂ progeny exhibited a continuous variation for tembotrione injury and recovery. Therefore, to perform a χ^2 -test frequency of segregation of tembotrione resistance or susceptibility in F₂ progeny, the plants that had more than 80% tembotrione injury were grouped as susceptible and others as resistant. The observed segregation of resistant:susceptible (R:S) ratios from both the crosses did not comply with the expected ratios of 3:1 (R:S) for a single gene inherited trait, indicating that more than one gene is involved in tembotrione resistance in G-200 or G-350 genotypes of sorghum (Table 4).

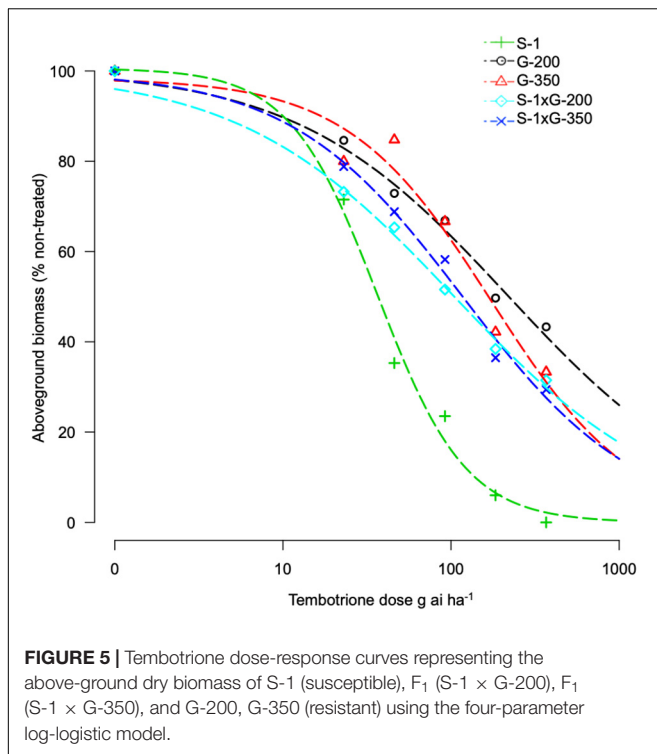
Mapping Tembotrione Resistance

To map the genomic loci controlling tembotrione resistance, a total of 208,376 SNPs were obtained using GBS from 150 F₂ progeny (S-1 × G-200) and parents (S-1, G-200). A subset of 1,954 SNP markers polymorphic to both parents with less than 30% missing values were retained. Further, filtering for missing rate (> 90%), strong segregation distortion, marker distribution, and redundant markers resulted in a total of 696 markers that were used for construction of a linkage map. The map of 1,021 cM was prepared which had an average distance of 1.7 cM between two adjacent markers. A total of three QTLs on chromosomes 2, 4, and 8 were mapped with a high LOD score (LOD > 3) (Figure 7 and Table 5) for two traits, RE, i.e., difference between leaf chlorophyll index at 2 and 4 WAT, and visual scoring at 2 WAT VI obtained from 150 F₂ plants (Supplementary Figure S1). The LOD score of detected QTLs ranged from 3.0 to 6.0 and the phenotypic variations explained (PVE) values ranged from 9 to 44%.

DISCUSSION

Identification and Characterization of Tembotrione-Resistant Sorghum

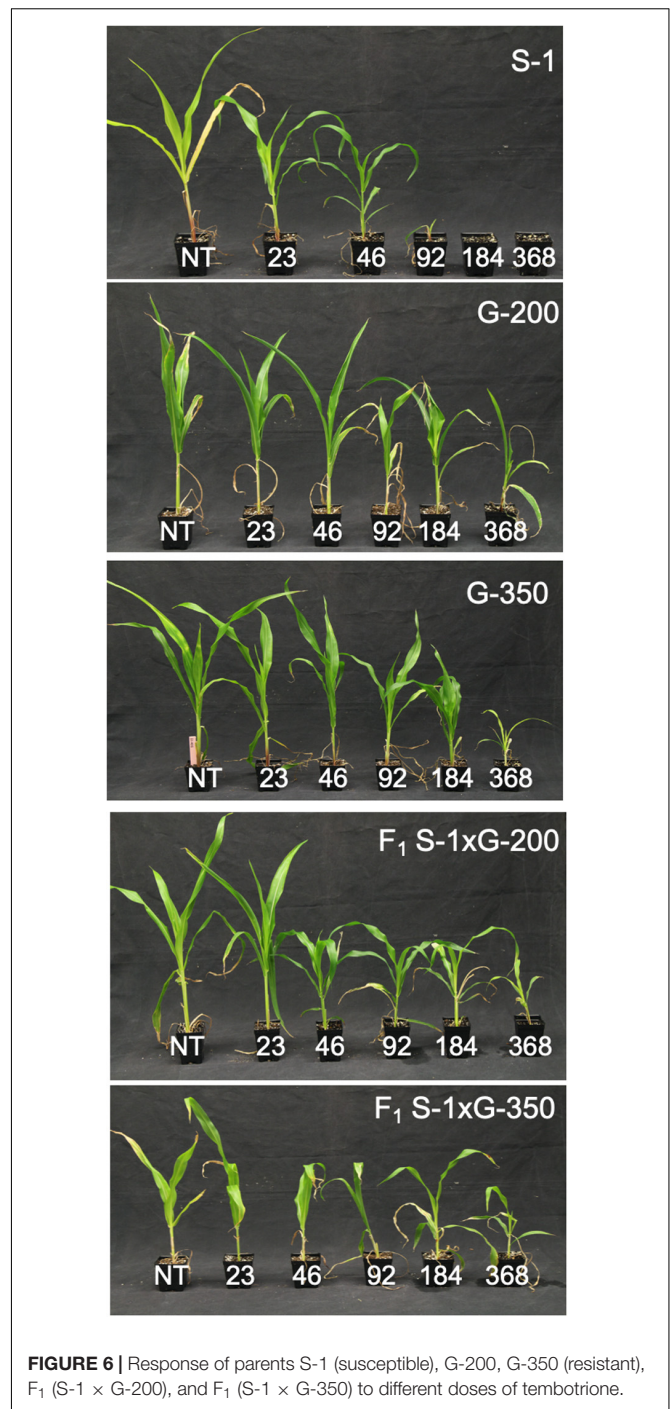
Tembotrione, an HPPD inhibitor, is widely used in corn for POST control of both grasses and broad-leaved weeds, but not



registered for use in sorghum due to crop injury (Damalas et al., 2018). In this study, we identified two sorghum genotypes, G-200 and G-350, with a high level of resistance and one genotype S-1 susceptible to tembotrione from SAP. Even though cultivated sorghum hybrids are susceptible to tembotrione, variable levels of responses were reported in sorghum germplasm upon treatment with tembotrione (Cunha et al., 2016). The genotypes, G-200 and G-350, were found ~10-fold and 7-fold more resistant, respectively, under greenhouse and ~2-fold more resistant under field conditions as compared with S-1 and Pioneer 84G62 (Table 1). In susceptible genotypes, the application of tembotrione causes foliar bleaching and leaf necrosis followed by complete plant death (Galon et al., 2016). Interaction of genotypes with environmental factors such as rainfall, soil type, and weather conditions plays a key role in plant response to HPPD inhibitors (Bollman et al., 2008), which may explain the difference in the level of resistance between greenhouse and field conditions. Variation in the efficacy of HPPD inhibitors in response to environmental factors such as temperature and relative humidity were reported in several weeds (Johnson and Young, 2002; Godar et al., 2015). However, in both greenhouse and field conditions, G-200 and G-350 survived the field dose of tembotrione, whereas S-1 or Pioneer 84G62 were severely injured (Table 1).

Mechanism of Tembotrione Resistance

The natural resistance to tembotrione in G-200 and G-350 appears to be not conferred by any alteration to the molecular target of this herbicide, i.e., *HPPD* gene, because no difference in *HPPD* gene sequence was found between the resistant (G-200 and



G-300) or susceptible (S-1) genotype (Supplementary Data S1). Likewise, no naturally evolved mutations in the *HPPD* gene that confer resistance to HPPD inhibitors were found in plants (Lu et al., 2020). Nonetheless, recently, soybean varieties resistant to HPPD inhibitors were developed through transgenic technology by inserting an insensitive *HPPD* gene (Siehl et al., 2014; Dreesen et al., 2018).

The CYP enzymes are known to metabolize HPPD inhibitors such as mesotrione (Ma et al., 2013; Nakka et al., 2017),

TABLE 3 | Regression parameters describing the response of sorghum genotypes and their F₁ progeny to tembotrione under greenhouse conditions.

Genotype	GR ₅₀ (SE)	RI (R/S)
S-1	36 (4)	1.0
S-1 × G-200	104 (27)**	2.8
S-1 × G-350	117 (25)**	3.2
G-350	172 (33)**	4.0
G-200	218 (60)**	6.0

GR₅₀, dose required for 50% growth reduction; SE, standard error; RI, resistance index; R/S, resistant/susceptible. Significantly different from S-1 at ** $p \leq 0.01$.

TABLE 4 | Chi-square analysis of the segregation of tembotrione-resistant (R) and tembotrione-susceptible (S) phenotypes in sorghum F₂ progeny at 4 weeks after treatment.

Cross	Run	Total	R	S	P-value
S-1 × G-200	1	200	168	32	0.0032**
	2	150	124	26	0.0372*
	Runs 1 and 2 combined	350	292	58	0.0003**
S-1 × G-350	1	220	197	23	0.0000**

Significantly different at ** $p \leq 0.01$, * $p \leq 0.05$.

tembotrione (Küpper et al., 2018), or topramezone (Elmore et al., 2015) in naturally evolved resistant weed biotypes. In this research, the genotypes G-200 and G-350 exhibited significant biomass reduction in response to pre-treatment with CYP inhibitors, malathion, or PBO followed by tembotrione application, suggesting that tembotrione is metabolized by CYP activity (Figures 3, 4). Similarly, the use of these inhibitors, followed by tembotrione, showed ~10% more biomass reduction compared with tembotrione alone treatments in a tembotrione-resistant common waterhemp population (Oliveira et al., 2018). Metabolism of tembotrione by hydroxylation followed by glycosylation, catalyzed by CYPs, has also been reported in a tembotrione-resistant Palmer amaranth biotype (Küpper et al., 2018). Furthermore, RNA-Seq analysis revealed differential expression of several CYP genes, for example, three to fourfold upregulation of *CYP72A219* and *CYP81E8*, respectively, was found in the same aforementioned tembotrione-resistant Palmer amaranth biotype (Küpper, 2018). In corn, multiple CYPs located in *nsf1* locus were found to metabolize tembotrione, mesotrione, as well as other herbicides (Williams and Pataky, 2010).

Inheritance of Tembotrione Resistance in Sorghum

Based on the response of F₁ progeny (S-1 × G-200 and S-1 × G-350) to tembotrione treatment, we found that the tembotrione resistance in G-200 and G-350 is a partially dominant trait (Figures 5, 6 and Table 3). Furthermore, F₂ data demonstrated that this resistance is controlled by multiple genes (Table 4). Genetic analyses of sweet corn inbred lines revealed a single recessive allele controlling the sensitivity to tembotrione (Williams and Pataky, 2008). The genetic basis of tembotrione resistance is not extensively studied in plants;

however, mesotrione (another widely used HPPD inhibitor) resistance in several common waterhemp populations across US Midwest (Huffman et al., 2015; Kohlhasse et al., 2018; Oliveira et al., 2018) was found to be inherited by a partially dominant polygenic trait.

QTL Mapping

We mapped three QTLs associated with tembotrione resistance on chromosomes 2, 4, and 8 using the sequence data from 150 F₂ plants from S-1 × G-200 cross (Figure 7). To our knowledge, this is the first report of QTLs associated with tembotrione resistance in grain sorghum. The QTL mapped using RE were previously reported for other traits in sorghum related to chlorophyll fluorescence (Fiedler et al., 2014) and photochemical quenching (Ortiz et al., 2017; Table 5); the QTLs mapped on chromosomes 2 and 4 using VI were novel QTLs and not previously reported for any other trait. The QTLs need to be tested in multiple environments with more number of F₂ plants and markers to improve the estimation accuracy, and experiments are in progress to further fine map and identify the precise location of the gene(s) responsible for tembotrione resistance in grain sorghum.

As mentioned earlier, our data indicate that the tembotrione resistance is a polygenic trait, and such traits can express differently in different genetic backgrounds. Therefore, tembotrione resistance can potentially be improved by crossing G-200 and G-350 or with other commercial genetic backgrounds. Such work has been reported to enhance the performance of quantitative traits in different genetic backgrounds and environmental conditions such as drought (Reddy et al., 2009), stay green (Subudhi et al., 2000), cold tolerance (Knoll and Ejeta, 2008), and yield (Nagaraja-Reddy et al., 2013) in grain sorghum. Therefore, there is enormous potential for improving tembotrione resistance by testing the expression of this trait in different genetic backgrounds and for the development of tembotrione-resistant sorghum varieties.

Because sorghum can outcross with closely related wild and weedy species, such as johnsongrass or shattercane, one of the major concerns of the development of herbicide-resistant sorghum varieties has been a natural transfer of such resistance into these weed species (Ohadi et al., 2017). However, recent reports suggest that the outcrossing rate of sorghum with johnsongrass was as low as ~1% under controlled conditions (Hodnett et al., 2019) and 2–16% with shattercane under field conditions (Schmidt et al., 2013). Although the possibility of outcrossing is minimal, if an herbicide resistance trait escapes into the wild species, necessary stewardship practices must be developed and integrated into sorghum weed management practices.

In conclusion, we have identified sorghum genotypes (G-200 and G-350) with natural resistance to tembotrione from the SAP, which can potentially be used to introgress the tembotrione resistance into breeding lines by conventional or marker-assisted breeding methods. CYP-inhibitor assay suggested CYP-mediated metabolism of tembotrione in the resistant genotypes. Genetic

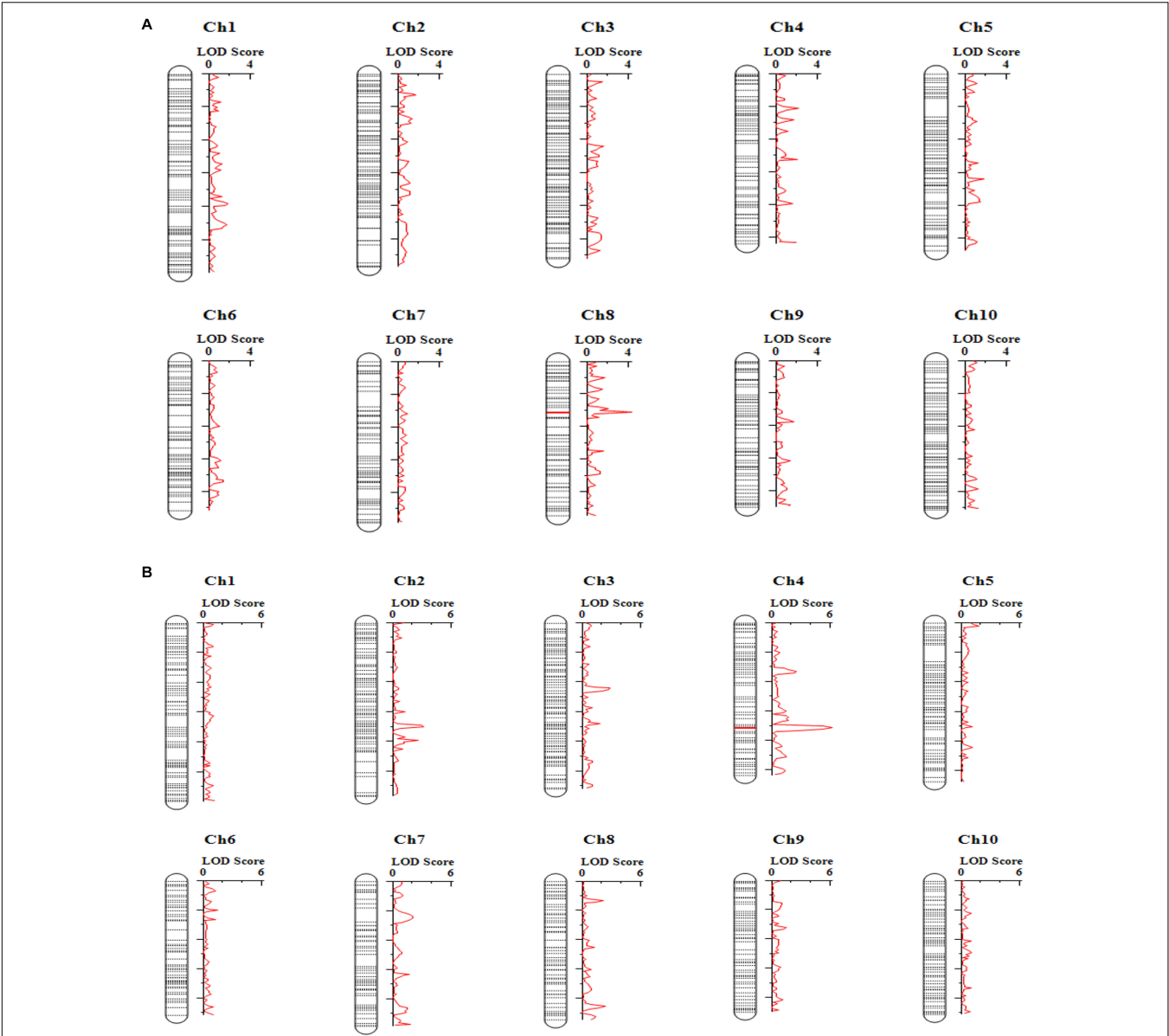


FIGURE 7 | Quantitative trait loci (QTLs) detected from analysis of 150 plants from a single F₂ family of S-1 × G-200 for different traits: **(A)** recovery (RE); **(B)** visual injury 21 days after treatment (VI).

TABLE 5 | Quantitative trait loci (QTLs) detected for the recovery (RE) and visual injury 3 weeks after treatment (VI) along with logarithm of the odds (LOD) and phenotypic variation explained (PVE) explained by QTLs.

Trait	QTL	Left marker	Right marker	LOD	PVE (%)	Add	Previously reported trait	References
RE	qRE8.1	Ch8_20217087	Ch8_20519857	4.32	44.35	−1.97	Efficiency of PSII reaction centers, chlorophyll fluorescence	Fiedler et al., 2014; Ortiz et al., 2017
VI	qVI4.0	Ch4_46023941	Ch4_48400958	6.20	21.23	17.18	–	–
VI	qVI2.1	Ch2_45821038	Ch2_46847104	3.15	9.71	11.79	–	–

analyses of F₁ and F₂ progeny demonstrated that the resistance is a partially dominant polygenic trait. Furthermore, GBS-based QTL mapping revealed three QTLs associated with tembotrione resistance in grain sorghum. Future research needs to be focused on incorporating the resistant trait with elite breeding varieties, testing the hybrid performance, and improving herbicide resistance in high yielding and stress tolerance hybrids.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in NCBI (BioProject ID: PRJNA657005).

AUTHOR CONTRIBUTIONS

BP performed the inheritance experiments, CYP-inhibitor assays, QTL mapping, curated the data, and wrote the first draft of the manuscript. AV performed *in vitro* and greenhouse experiments. BP, AV, and ARV conducted dose-response experiments and field testing. MT assisted high-throughput DNA extraction. MZ prepared the GBS library. SL, RS, GL, and BP performed the

QTL mapping. TT provided the seed material and input on genetic analyses. MJ and PVVP designed and conceived the experiments, analyzed the data, and helped to draft the final manuscript. All authors read, edited, and approved the final version of the manuscript.

FUNDING

The authors thank funding support from the Center for Sorghum Improvement, the Department of Agronomy and the College of Agriculture, Kansas State University. Graduate Student Assistantship to Pandian from Kansas Grain Sorghum Commission is highly appreciated. This is contribution number 20-348-J from the Kansas Agricultural Experiment Station.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.596581/full#supplementary-material>

REFERENCES

- Abit, M. J. M., Al-Khatib, K., Olson, B. L., Stahlman, P. W., Geier, P. W., Thompson, C. R., et al. (2011). Efficacy of postemergence herbicides tankmixes in aryloxyphenoxypionate-resistant grain sorghum. *Crop Prot.* 30(12), 1623–1628. doi: 10.1016/j.cropro.2011.08.011
- Abit, M. J. M., Al-Khatib, K., Regehr, D. L., and Tuinstra, M. R. (2009). Differential response of grain sorghum hybrids to foliar-applied mesotrione. *Weed Technol.* 23, 28–33. doi: 10.1614/WT-08-086.1
- Ahrens, H., Lange, G., Müller, T., Rosinger, C., Willms, L., and van Almsick, A. (2013). 4-Hydroxyphenylpyruvate dioxygenase inhibitors in combination with safeners: solutions for modern and sustainable agriculture. *Angew. Chem. Int. Ed.* 52, 9388–9398. doi: 10.1002/anie.201302365
- Bai, G., Kolb, F. L., Shaner, G., and Domier, L. L. (1999). Amplified fragment length polymorphism markers linked to a major quantitative trait locus controlling scab resistance in wheat. *Phytopathology* 89, 343–348. doi: 10.1094/PHYTO.1999.89.4.343
- Bollman, J. D., Boerboom, C. M., Becker, R. L., and Fritz, V. A. (2008). Efficacy and tolerance to HPPD-Inhibiting herbicides in sweet corn. *Weed Technol.* 22, 666–674. doi: 10.1614/WT-08-036.1
- Busi, R., Gaines, T. A., and Powles, S. (2017). Phorate can reverse P450 metabolism-based herbicide resistance in *Lolium rigidum*. *Pest Manag. Sci.* 73, 410–417. doi: 10.1002/ps.4441
- Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., et al. (2008). Community resources and strategies for association mapping in sorghum. *Crop Sci.* 48, 30–40. doi: 10.2135/cropsci2007.02.0080
- Churchill, G. A., and Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 3, 963–971.
- Ciampitti, I. A., and Prasad, P. V. V. (eds). (2019). *Sorghum: State of the Art and Future Perspectives*. Agronomy Monograph, Vol. 58. Madison, WI: American Society of Agronomy.
- Cochran, W. G. (1952). The χ^2 test of goodness of fit. *Ann. Math. Stat.* 23, 315–345.
- Cunha, F. N., Vidal, V. M., Silva, N. F., Soares, F. A. L., Dahlberg, J., Berenji, J., et al. (2016). Selectivity of the tembotrione herbicide to sorghum. *Rev. Bras.* 15, 281–293. doi: 10.18512/1980-6477/rbms.v15n2p281-293
- Dahlberg, J., Berenji, J., Sikora, V., and Latković, D. (2012). *Assessing Sorghum [Sorghum bicolor (L) Moench] Germplasm for New Traits: Food, Fuels and Unique Uses*. Maydica 56. Available online at: <https://journals-crea.4science.it/index.php/maydica/article/view/688> (accessed May 29, 2020).
- Damalas, C. A., Gitsopoulos, T. K., Koutroubas, S. D., Alexoudis, C., and Georgoulas, I. (2018). Weed control and selectivity in maize (*Zea mays* L.) with tembotrione mixtures. *Int. J. Pest Manage.* 64, 11–18. doi: 10.1080/09670874.2017.1293307
- Dankov, K., Busheva, M., Stefanov, D., and Apostolova, E. L. (2009). Relationship between the degree of carotenoid depletion and function of the photosynthetic apparatus. *J. Photochem. Photobiol. B* 96, 49–56. doi: 10.1016/j.jphotobiol.2009.04.004
- de Mendiburu, F. (2014). *Agricolae: Statistical Procedures for Agricultural Research. R Package Version 1*. Available online at: <https://rdrr.io/cran/agricolae/f/inst/doc/tutorial.pdf> (accessed June 15, 2020).
- Development R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Available online at: <http://finzi.psych.upenn.edu/R/library/dplR/doc/intro-dplR.pdf> (accessed June 15, 2020).
- Dille, J. A., Stahlman, P. W., Thompson, C. R., Bean, B. W., Soltani, N., and Sikkema, P. H. (2020). Potential yield loss in grain sorghum (*Sorghum bicolor*) with weed interference in the United States. *Weed Technol.* 34, 624–629. doi: 10.1017/wet.2020.12
- Dreesen, R., Capt, A., Oberdoerfer, R., Coats, I., and Kenneth, E. P. (2018). Characterization and safety evaluation of HPPD W336, a modified 4-hydroxyphenylpyruvate dioxygenase protein, and the impact of its expression on plant metabolism in herbicide-tolerant MST-FG072-2 soybean. *Regul. Toxicol. Pharmacol.* 97, 170–185. doi: 10.1016/j.yrtph.2018.06.002
- Elmore, M. T., Brosnan, J. T., Armel, G. R., Kopsell, D. A., Best, M. D., Mueller, T. C., et al. (2015). Cytochrome P450 inhibitors reduce creeping bentgrass (*Agrostis stolonifera*) tolerance to topramezone. *PLoS One* 10:e0130947. doi: 10.1371/journal.pone.0130947
- Fiedler, K., Bekele, W. A., Duensing, R., Gründig, S., Snowden, R., Stützel, H., et al. (2014). Genetic dissection of temperature-dependent sorghum growth during juvenile development. *Theor. Appl. Genet.* 127, 1935–1948. doi: 10.1007/s00122-014-2350-7
- Gaines, T. A., Duke, S. O., Morran, S., Rigon, C. A. G., Tranel, P. J., Küpper, A., et al. (2020). Mechanisms of evolved herbicide resistance. *J. Biol. Chem.* 295, 10307–10330. doi: 10.1074/jbc.REV120.013572
- Galon, L., Fernandes, F. F., Andres, A., da Silva, A. F., and Forte, C. T. (2016). Selectivity and efficiency of herbicides in weed control on sweet sorghum. *Pesqui. Agropec. Trop.* 46, 123–131. doi: 10.1590/1983-40632016v4639431
- Godar, A. S., Varanasi, V. K., Nakka, S., Prasad, P. V. V., Thompson, C. R., and Mithila, J. (2015). Physiological and molecular mechanisms of differential

- sensitivity of palmer amaranth (*Amaranthus palmeri*) to mesotrione at varying growth temperatures. *PLoS One* 10:e0126731. doi: 10.1371/journal.pone.0126731
- Hawkes, T. R., Langford, M. P., Viner, R. C., Vernooij, B. T. M., and Dale, R. (2016). *Mutant Hydroxyphenylpyruvate Dioxygenase Polypeptides and Methods of Use*. US Patent. Available online at: <https://patentimages.storage.googleapis.com/4c/72/11/9e612b22dae2f9/US9388393.pdf> (accessed May 30, 2020).
- Heap, I. (2020). *International Herbicide-Resistant Weed Database*. Available online at: www.weedscience.org (accessed November 24, 2020).
- Hodnett, G. L., Ohadi, S., Pugh, N. A., Bagavathiannan, M. V., and Rooney, W. L. (2019). Sorghum bicolor x *S. halepense* interspecific hybridization is influenced by the frequency of 2n gametes in *S. bicolor*. *Sci. Rep.* 9:17901. doi: 10.1038/s41598-019-53193-3
- Huffman, J., Hausman, N. E., Hager, A. G., Riechers, D. E., and Tranel, P. J. (2015). Genetics and inheritance of nontarget-site resistances to atrazine and mesotrione in a waterhemp (*Amaranthus tuberculatus*) population from Illinois. *Weed Sci.* 63, 799–809. doi: 10.1614/WS-D-15-00055.1
- Johnson, B. C., and Young, B. G. (2002). Influence of temperature and relative humidity on the foliar activity of mesotrione. *Weed Sci.* 50, 157–161.
- Knezevic, S. Z., Streibig, J. C., and Ritz, C. (2007). Utilizing R software package for dose-response studies: the concept and data analysis. *Weed Technol.* 21, 840–848. doi: 10.1614/WT-06-161.1
- Knoll, J., and Ejeta, G. (2008). Marker-assisted selection for early-season cold tolerance in sorghum: QTL validation across populations and environments. *Theor. Appl. Genet.* 116, 541–553. doi: 10.1007/s00122-007-0689-8
- Kohlhase, D. R., Edwards, J. W., and Owen, M. D. K. (2018). Inheritance of 4-hydroxyphenylpyruvate dioxygenase inhibitor herbicide resistance in an *Amaranthus tuberculatus* population from Iowa, USA. *Plant Sci.* 274, 360–368. doi: 10.1016/j.plantsci.2018.06.004
- Küpper, A. (2018). *Molecular Genetics of Herbicide Resistance in Palmer Amaranth (Amaranthus palmeri): Metabolic Tembotrione Resistance and Geographic Origin of Glyphosate Resistance*. Ph.D. dissertation, Colorado State University, Fort Collins, CO.
- Küpper, A., Peter, F., Zöllner, P., Lorentz, L., Tranel, P. J., Beffa, R., et al. (2018). Tembotrione detoxification in 4-hydroxyphenylpyruvate dioxygenase (HPPD) inhibitor-resistant Palmer amaranth (*Amaranthus palmeri* S. Wats.). *Pest Manag. Sci.* 74, 2325–2334. doi: 10.1002/ps.4786
- Lee, D. L., Knudsen, C. G., Michaely, W. J., Chin, H.-L., Nguyen, N. H., Carter, C. G., et al. (1998). The structure-activity relationships of the triketone class of HPPD herbicides. *Pestic. Sci.* 54, 377–384. doi: 10.1021/bk-2001-0774.ch002
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Lu, H., Yu, Q., Han, H., Owen, M. J., and Powles, S. B. (2020). Evolution of resistance to HPPD-inhibiting herbicides in a wild radish population via enhanced herbicide metabolism. *Pest Manag. Sci.* 76, 1929–1937. doi: 10.1002/ps.5725
- Ma, R., Kaundun, S. S., Tranel, P. J., Riggins, C. W., McGinness, D. L., Hager, A. G., et al. (2013). Distinct detoxification mechanisms confer resistance to mesotrione and atrazine in a population of waterhemp. *Plant Physiol.* 163, 363–377. doi: 10.1104/pp.113.223156
- Mace, E., Innes, D., Hunt, C., Wang, X., Tao, Y., Baxter, J., et al. (2019). The Sorghum QTL Atlas: a powerful tool for trait dissection, comparative genomics and crop improvement. *Theor. Appl. Genet.* 132, 751–766. doi: 10.1007/s00122-018-3212-5
- Matringe, M., Sailland, A., Pelissier, B., Rolland, A., and Zink, O. (2005). p-Hydroxyphenylpyruvate dioxygenase inhibitor-resistant plants. *Pest Manag. Sci. Formerly Pestic. Sci.* 61, 269–276. doi: 10.1002/ps.997
- McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., et al. (2017). The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* 93, 338–354. doi: 10.1111/tpj.13781
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001
- Nagaraja-Reddy, R., Madhusudhana, R., Murali Mohan, S., Chakravarthi, D. V. N., Mehtre, S. P., Seetharama, N., et al. (2013). Mapping QTL for grain yield and other agronomic traits in post-rainy sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.* 126, 1921–1939. doi: 10.1007/s00122-013-2107-8
- Nakka, S., Godar, A. S., Wani, P. S., Thompson, C. R., Peterson, D. E., Roelofs, J., et al. (2017). Physiological and molecular characterization of hydroxyphenylpyruvate dioxygenase (HPPD)-inhibitor resistance in Palmer amaranth (*Amaranthus palmeri* S. Wats.). *Front. Plant Sci.* 8:555. doi: 10.3389/fpls.2017.00555
- Ohadi, S., Hodnett, G., Rooney, W., and Bagavathiannan, M. (2017). Gene flow and its consequences in sorghum spp. *CRC Crit. Rev. Plant Sci.* 36, 367–385. doi: 10.1080/07352689.2018.1446813
- Oliveira, M. C., Gaines, T. A., Dayan, F. E., Patterson, E. L., Jhala, A. J., and Knezevic, S. Z. (2018). Reversing resistance to tembotrione in an *Amaranthus tuberculatus* (var. rudis) population from Nebraska, USA with cytochrome P450 inhibitors. *Pest Manag. Sci.* 74, 2296–2305. doi: 10.1002/ps.4697
- Ortiz, D., Hu, J., and Salas Fernandez, M. G. (2017). Genetic architecture of photosynthesis in *Sorghum bicolor* under non-stress and cold stress conditions. *J. Exp. Bot.* 68, 4545–4557. doi: 10.1093/jxb/erx276
- Ott, A., Liu, S., Schnable, J. C., ‘Eddy’ Yeh, C.-T., Wang, K.-S., and Schnable, P. S. (2017). tGBS® genotyping-by-sequencing enables reliable genotyping of heterozygous loci. *Nucleic Acids Res.* 45:e178. doi: 10.1093/nar/gkx853
- Pandian, B. A., Sathishraj, R., Djanaguiraman, M., Prasad, P. V. V., and Jugulam, M. (2020). Role of cytochrome P450 enzymes in plant stress response. *Antioxidants* 9:454. doi: 10.3390/antiox9050454
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457, 551–556. doi: 10.1038/nature07723
- Rakshit, S., and Bellundagi, A. (2019). “Chapter 5 – conventional breeding techniques in sorghum,” in *Breeding Sorghum for Diverse End Uses*, eds C. Aruna, K. B. R. S. Visarada, B. V. Bhat, and V. A. Tonapi (Cambridge: Woodhead Publishing), 77–91.
- Reddy, B. V. S., Ramesh, S., Reddy, P. S., and Kumar, A. A. (2009). Genetic enhancement for drought tolerance in sorghum. *Plant Breed. Rev.* 31:189. doi: 10.1002/9780470593783.ch3
- Ritz, C., and Streibig, J. C. (2005). Bioassay analysis using R. *J. Stat. Softw.* 12, 1–22.
- Roozeboom, K. L., and Prasad, P. V., (2019). “Sorghum growth and development,” in *Sorghum: State of the Art and Future Perspectives Agronomy Monographs*, eds I. A. Ciampitti and P. V. V. Prasad (Madison, WI: American Society of Agronomy and Crop Science Society of America, Inc).
- Schmidt, J. J., Pedersen, J. F., Bernards, M. L., and Lindquist, J. L. (2013). Rate of shattercane × sorghum hybridization in situ. *Crop Sci.* 53, 1677–1685.
- Shyam, C., Jhala, A. J., Kruger, G., and Jugulam, M. (2019). Rapid metabolism increases the level of 2,4-D resistance at high temperature in common waterhemp (*Amaranthus tuberculatus*). *Sci. Rep.* 9:16695. doi: 10.1038/s41598-019-53164-8
- Siehl, D. L., Tao, Y., Albert, H., Dong, Y., Heckert, M., Madrigal, A., et al. (2014). Broad 4-hydroxyphenylpyruvate dioxygenase inhibitor herbicide tolerance in soybean with an optimized enzyme and expression cassette. *Plant Physiol.* 166, 1162–1176. doi: 10.1104/pp.114.247205
- Siminszky, B. (2006). Plant cytochrome P450-mediated herbicide metabolism. *Phytochem. Rev.* 5, 445–458. doi: 10.1007/s11101-006-9011-7
- Smith, K., Scott, B., Espinoza, L., and Kelley, J. (2010). *Weed Control in Grain Sorghum. Grain Sorghum Production Handbook*. Little Rock, AR: Cooperative Extension Service, University of Arkansas, 47–49.
- Staggenborg, S. A., Dhuyvetter, K. C., and Gordon, W. B. (2008). Grain sorghum and corn comparisons: yield, economic, and environmental responses. *Agron. J.* 100, 1600–1604. doi: 10.2134/agronj2008.0129
- Stahlman, P. W., Wicks, G. A., Smith, C. W., and Frederiksen, R. A. (2000). *Weeds and Their Control in Grain Sorghum. Sorghum: Origin, History, Technology, and Production*. New York, NY: Wiley, 535–590.
- Subudhi, P. K., Rosenow, D. T., and Nguyen, H. T. (2000). Quantitative trait loci for the stay green trait in sorghum (*Sorghum bicolor* L. Moench): consistency across genetic backgrounds and environments. *Theor. Appl. Genet.* 101, 733–741. doi: 10.1007/s001220051538

- Taylor, J. R. N., Schober, T. J., and Bean, S. R. (2006). Novel food and non-food uses for sorghum and millets. *J. Cereal Sci.* 44, 252–271. doi: 10.1016/j.jcs.2006.06.009
- Thompson, C. R., Dille, J. A., and Peterson, D. E. (2019). “Weed competition and management in sorghum,” in *Sorghum: State of the Art and Future Perspectives* Agronomy Monographs, eds I. A. Ciampitti and P. V. V. Prasad (Madison, WI: American Society of Agronomy and Crop Science Society of America, Inc).
- USDA-NASS (2020). Available online at: <https://www.nass.usda.gov/> (accessed 12 May 2020).
- Valadabad, A., Mazaheri, D., Nour-Mohamad, G., and Hashemi, D. S. (2000). Performance of the effect of drought stress on qualitative and quantitative characters of corn, sorghum and millet. *Iran. J. Crop Sci.* 2, 39–47.
- van Almsick, A. (2009). New HPPD-inhibitors—a proven mode of action as a new hope to solve current weed problems. *Outlooks Pest Manag.* 20, 27–30. doi: 10.1564/20feb09
- Wickham, H., and Wickham, M. H. (2007). *The ggplot Package*. Available online at: <http://ftp.uni-bayreuth.de/math/statlib/R/CRAN/doc/packages/ggplot.pdf> (accessed July 6, 2020).
- Williams, M. M., and Pataky, J. K. (2008). Genetic basis of sensitivity in sweet corn to tembotrione. *Weed Sci.* 56, 364–370. doi: 10.1614/WS-D-09-00058.1
- Williams, M. M., and Pataky, J. K. (2010). Factors affecting differential sensitivity of sweet corn to HPPD-inhibiting herbicides. *Weed Sci.* 58, 289–294. doi: 10.1614/WS-D-09-00058.1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Pandian, Varanasi, Vennapusa, Sathishraj, Lin, Zhao, Tunnell, Tesso, Liu, Prasad and Jugulam. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



OPEN ACCESS

Edited by:

Velu Govindan,
International Maize and Wheat
Improvement Center, Mexico

Reviewed by:

Deepmala Sehgal,
International Maize and Wheat
Improvement Center, Mexico
Dinakaran Elango,
Pennsylvania State University (PSU),
United States

*Correspondence:

Vânia C. R. Azevedo
v.azevedo@cgiar.org
Mani Vetriventhan
m.vetriventhan@cgiar.org

†ORCID:

Victor Allan
orcid.org/0000-0001-7230-081X
Mani Vetriventhan
orcid.org/0000-0001-9111-2320
Ramachandran Senthil
orcid.org/0000-0003-4017-8180
S. Geetha
orcid.org/0000-0002-1403-4741
Santosh Deshpande
orcid.org/0000-0002-7182-3456
Abhishek Rathore
orcid.org/0000-0001-6887-4095
Vinod Kumar
orcid.org/0000-0001-7837-7770
Prabhat Singh
orcid.org/0000-0001-7827-697X
Surender Reddymalla
orcid.org/0000-0002-0983-2474
Vânia C. R. Azevedo
orcid.org/0000-0002-9088-078X

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 26 July 2020

Accepted: 05 November 2020

Published: 14 December 2020

Genome-Wide DArTSeq Genotyping and Phenotypic Based Assessment of Within and Among Accessions Diversity and Effective Sample Size in the Diverse Sorghum, Pearl Millet, and Pigeonpea Landraces

Victor Allan^{1†}, Mani Vetriventhan^{2*†}, Ramachandran Senthil^{2†}, S. Geetha^{1†}, Santosh Deshpande^{2†}, Abhishek Rathore^{2†}, Vinod Kumar^{2†}, Prabhat Singh^{2†}, Surender Reddymalla^{2†} and Vânia C. R. Azevedo^{2*†}

¹ Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University (TNAU), Coimbatore, India, ² International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Germplasm should be conserved in such a way that the genetic integrity of a given accession is maintained. In most genebanks, landraces constitute a major portion of collections, wherein the extent of genetic diversity within and among landraces of crops vary depending on the extent of outcrossing and selection intensity infused by farmers. In this study, we assessed the level of diversity within and among 108 diverse landraces and wild accessions using both phenotypic and genotypic characterization. This included 36 accessions in each of sorghum, pearl millet, and pigeonpea, conserved at ICRISAT genebank. We genotyped about 15 to 25 individuals within each accession, totaling 1,980 individuals using the DArTSeq approach. This resulted in 45,249, 19,052, and 8,211 high-quality single nucleotide polymorphisms (SNPs) in pearl millet, sorghum, and pigeonpea, respectively. Sorghum had the lowest average phenotypic (0.090) and genotypic (0.135) within accession distances, while pearl millet had the highest average phenotypic (0.227) and genotypic (0.245) distances. Pigeonpea had an average of 0.203 phenotypic and 0.168 genotypic within accession distances. Analysis of molecular variance also confirms the lowest variability within accessions of sorghum (26.3%) and the highest of 80.2% in pearl millet, while an intermediate in pigeonpea (57.0%). The effective sample size required to capture maximum variability and to retain rare alleles while regeneration ranged from 47 to 101 for sorghum, 155 to 203 for pearl millet, and 77 to 89 for pigeonpea accessions. This study will support genebank curators, in understanding the dynamics of population within and among accessions, in devising appropriate germplasm conservation strategies, and aid in their utilization for crop improvement.

Keywords: DArTseq, within accession diversity, effective population size, landraces, pearl millet, pigeonpea, regeneration, sorghum

INTRODUCTION

Plant genetic resources include landraces, wild and weedy relatives, improved cultivars, etc. which are of potential value as a resource for present and future generations of people. Landraces occupy a major portion in collections conserved in genebanks. Landraces possess a multifaceted evolutionary history and a vast diversity, primarily associated with humans, also influenced by both natural and farmers' informal selections (Hawkes, 1983). The high variability and genetic diversity of landraces are well-known. Harlan (1965) reported the gene-flow from weeds to landraces and several other authors (Ellstrand et al., 1999; Jarvis and Hodgkin, 1999; Messeguer, 2003; Gompert and Buerkle, 2016) reported the transfer and diffusion of genes into landraces from various sources in both self and out-crossing species. Harlan (1971) emphasized landraces as genetically dynamic populations, and a result of millennia of artificial and natural selection, also Hawkes (1983) described landraces as highly diverse populations or a mixture of heterogeneous genotypes, and several other authors proposed various definitions to landraces, explaining their heterogeneity and genetic nature (Brown, 1978; Martin and Adams, 1987; Astley, 1991; Michaelis et al., 1991). Brown (1978), Bellon (2009), and Frankel and Soule (1981) explained the occurrence of within and between population genetic variation in landrace populations and further explained the within-population diversity is mainly an effect of heterogeneity over space and time. Many pieces of literature are available emphasizing the high variability in landraces, however, only a few studies are available investigating diversity within individuals of landrace accessions that are conserved in genebanks (Busso et al., 2000; Bhattacharjee et al., 2002), while few other studies focused on diversity within landrace populations conserved *in situ* (Djè et al., 1999; Pressoir and Berthaud, 2004; Dreisigacker et al., 2005; Al Khanjari et al., 2007; Jones et al., 2008; Hagenblad et al., 2012; Kyratzis et al., 2019).

Therefore, understanding the diversity within landraces is essential to make sure that, in genebanks the genetic integrity of a given accession is maintained with its innate variability and diversity without losing any rare allele variants. The major cause for allele loss in genebank accessions is genetic drift when accessions are regenerated with small sample sizes (Crossa, 1995). Mode of pollination being the key factor governing the frequencies of alleles within different individuals of a population, it influences the variability, quantum of diversity, gene flow and population dynamics behind evolution. Hammer et al. (1996) explained the effect of mode of pollination on genetic erosion of landraces, Zeven (1998) explained the attainment of gradual homozygosity within inbreeding landrace populations, and Villa et al. (2005) explained the influence of mode of reproduction in alteration of genetic structure of landraces. Genebanks exercise various scientific strategies to preserve the inherent genetic variability within each accession with theoretical foundations of various population genetic considerations, mainly the mode of reproduction, allelic frequencies, distribution of allelic variations, the proportion of rare alleles, etc. to maintain the genetic integrity of an accession. Rare alleles, however, are easily susceptible to random genetic drifts and can be lost permanently (Ramanatha

Rao and Hodgkin, 2002) when handled with inadequate scientific knowledge about the underlying population dynamics. Thus, appropriate conservation strategies with statistically estimated population sizes should be followed. In this study, we have chosen three crops that differ in pollination behavior, including highly cross pollinated pearl millet (>85%) (Burton, 1983), and often-cross pollinated sorghum (about 18%) (Barnaud et al., 2008) and pigeonpea (about 30%) (Saxena et al., 1990), to comparatively assess the within and between accession diversity. Landraces of these crops possess large variability within accessions, therefore chosen for this study.

Classical molecular markers used to assess the genetic diversity in these crops included SSR (Budak et al., 2003; Chandra-Shekara et al., 2007; Bashir et al., 2015), RFLP (Bhattacharjee et al., 2002; Govindaraj et al., 2009), ISSR (Kumar et al., 2006; Animasaun et al., 2015), RAPD (Chowdary et al., 1998; Chandra-Shekara et al., 2007), SRAP (Xie et al., 2010), etc. However, these molecular markers had constraints such as high cost of genotyping per sample and most of these technologies are gel-based and lacked the ability to rapidly analyze large number of marker loci. Recent technological developments in high throughput genotyping overcame these limitations and technologies like DArTSeq, by combining DArT (Diversity Array Technology) with NGS (Next Generation Sequencing), offered the flexibility of genome-wide characterization of germplasms, even without prior sequence information, parallelly providing a low-cost platform for high throughput marker genotyping. Several studies using DArTSeq on diversity and population structure assessments have been reported on various crops evidencing the potential scope of this technology in diversity assessment (Pailles et al., 2017; Raman et al., 2017; Barilli et al., 2018; Edet et al., 2018; Ndjioudjop et al., 2018). What makes DArTSeq to stand apart from other GBS (Genotyping By Sequencing) techniques is their method of complexity reduction that are targeted over the genomic coding regions and the additional advantage of genotyping without prior sequence information extends its scope even toward the under researched wild accessions. It also offers relatively better genome coverage with high reproducibility as DArTSeq is performed at higher sequencing depths and uses strict filtering criterions, it generates markers with less missing data compared to other GBS approaches.

With these background, this study aims (i) to assess genotypic (DArTSeq) and phenotypic characterization of geographically representative diverse sorghum, pearl millet, and pigeonpea landraces and wild accessions to comparatively investigate the extent of diversity within and among accessions, and (ii) to assess the minimum sample (population) size required to capture 95% of the alleles with an expected probability of 95%, from the least frequent allele or from the frequency of the rarest allele for each accession. The scope of this study aims to benefit genebank curators in understanding the dynamics of population within and among accessions, and devising proper sampling strategies (sample size) while regeneration, for effective genebank management and for their utilization in crop improvement. To the best of our knowledge, this study is the first of its kind, and no studies were found utilizing NGS for investigating within

accession diversity and sample size estimations, particularly for sorghum, pigeonpea, and pearl millet.

MATERIALS AND METHODS

Plant Material

This study investigated a total of 108 geographically diverse accessions of sorghum, pearl millet, and pigeonpea (**Supplementary Tables 1–3**) (**Figure 1**), conserved at ICRISAT genebank. Accessions of sorghum included 31 landraces and 5 wild accessions, collected from 26 different countries from 5 different continents, consisted of all the 5 basic races and all 10 intermediate races as classified by Harlan and de Wet (1972). Accessions of pearl millet consisted of 33 landraces and 3 wild accessions, collected from 19 different countries from 2 different continents, and accessions of pigeonpea included 36 landraces collected from 34 different countries from 5 different continents. All these 108 accessions were raised in fields during post-rainy 2018 at ICRISAT, Hyderabad, for phenotypic and genotypic characterization. Sorghum accessions were sown on black soil, whereas pearl millet and pigeonpea were sown on red soil. Accessions of sorghum occupied three-rows of 9 m length, spaced 75 cm between rows, with a plant-to-plant spacing of about 10 cm. Accessions of pearl millet were laid in 4-meter rows, with each accession occupying 4 rows, spaced 75 cm between

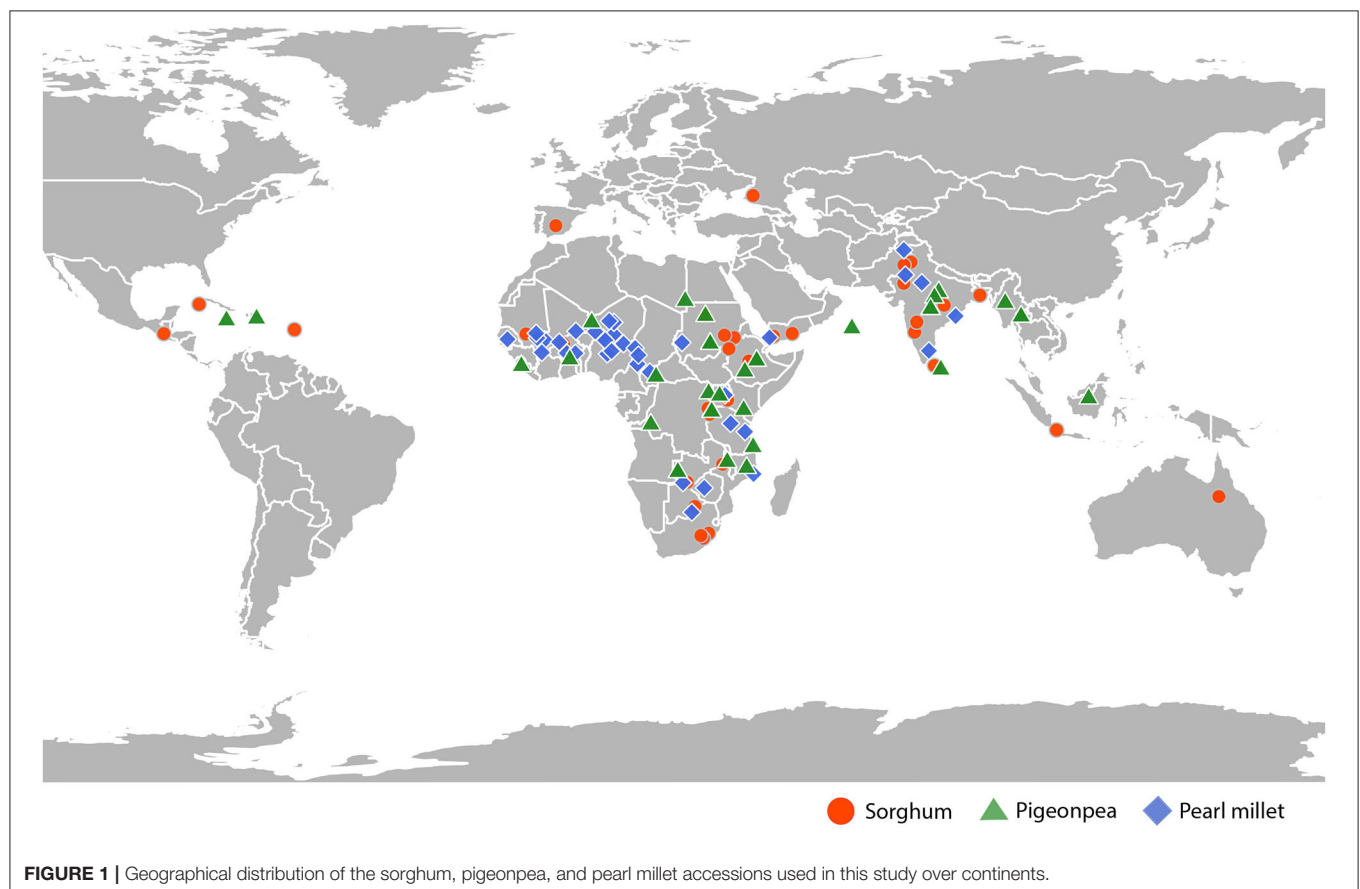
rows and 10 cm between plants. Each accession of pigeonpea occupied two rows of 9-meter length, spaced ~75 cm between rows and 50 cm between plants.

DNA Extraction, Complexity Reduction and Genotyping

Individual plants within each accession of sorghum, pearl millet, and pigeonpea were labeled with unique plant ID, and leaf samples were collected from 15 plants in each accession of sorghum and pigeonpea, and 25 plants from each accession of pearl millet, totaling a 540, 900 and 540 samples in sorghum, pearl millet, and pigeonpea respectively. Leaf samples were collected from 15 days old seedlings of sorghum and pearl millet, and 2-month old seedlings of pigeonpea. Collected leaf samples were sealed in zip lock bags or collected using the PCR plates with corresponding plant ID for each sample and packed with ice cubes, and sent for DNA extraction on the same day. The DNA extraction was carried out following the procedure reported by Mace et al. (2003) and the extracted genomic DNA samples were sent to DArT Private Limited in Canberra, Australia (www.diversityarrays.com) for DArTSeq genotyping.

SNP Filtering

The SNP markers from DArTSeq were filtered with a maximum threshold of 95% reproducibility, 80% call rate for markers, and 50% missing values over samples. The SNPs were not filtered for



minor allele frequencies (MAF) in order to preserve the rare allele variants, which have the main part of the focus in this study.

Phenotypic Evaluation

To capture maximum phenotypic variability, all the individual plants within each accession were labeled with unique plant ID and data on both qualitative and quantitative traits (Supplementary Table 4) were recorded for all the 3 crops following the respective crop descriptors (IBPGR and ICRISAT, 1993a,b,c), throughout the growing season. In this study, a large number of plants including those plants that were used for DArTSeq and also plants that were not sampled for DArTSeq were phenotyped. The total plant count for each accession ranged from 115 to 234 in sorghum, 51 to 116 in pearl millet. However, in pigeonpea, only 35 accessions had plant count over 10. Two accessions of pigeonpea had a plant count of <14, and remaining accessions had plant counts between 21 and 33. Thus, only data from the 35 accessions of pigeonpea was used for phenotypic analysis.

Phenotypic Data Analysis

Descriptive statistics such as mean, standard deviation, and standard error were computed for quantitative traits to assess the spread and distribution of the data. Preliminary analysis of phenotypic data included investigating diversity among accessions using the mean and range values. This was followed by *post-hoc* tests, which included Student Newman Keuls test (Newman, 1939; Keuls, 1952) and Levene's test (Levene, 1960) to verify statistical significance between means and homogeneity of variances, respectively. Gower distance metric (Gower, 1971) was used for within accession diversity assessment using both quantitative and qualitative data. Pairwise distances between individual plants were subjected to the ward.D2 agglomerative clustering algorithm (Murtagh and Legendre, 2014) with 100 bootstraps. The same set of analyses were applied to all three crops. R software v.3.6.0 (R Core Team, 2019) was used with R-CRAN packages like “cluster” (Maechler et al., 2019) for Gower's distance computation, “fpc” (Hennig, 2020) and “pvclust” (Suzuki et al., 2019) for bootstrapped clustering, “car” (Fox and Weisberg, 2019) and “agricolae” (de Mendiburu, 2013) for SNK test and Levene's test, respectively.

Genotypic Data Analysis

DArTSeq derived SNP data after filtering were used for analysis. Analysis of Molecular Variance (AMOVA) was computed as proposed by Excoffier et al. (1992), which partitioned the total variance into within and among population variance components. AMOVA was carried out considering each accession (with 15 or 25 individuals) as a separate population. For testing the significance, results of AMOVA were subjected to Monte Carlo's estimate of *p*-values with 99 permutations. Heterozygosity was estimated as reported by Nei (1973). For diversity assessment, Euclidean based modified Roger's distance metric (Goodman and Stuber, 1983) was used and distances between individual plants were computed, which was followed by ward.D2 agglomerative clustering (Murtagh and Legendre, 2014) and a dendrogram was produced. The “clusterboot”

function from the R-package “fpc” (Hennig, 2020) and the “aboot” function from the R-package “poppr” (Kamvar et al., 2015) were used to evaluate the clusters with 100 bootstraps. Shannon diversity (H') (Shannon, 1948) was calculated for each accession using the formula,

$$H' = - \sum_{a=1}^A p_i \log_2(p_i)$$

Where p_i is the estimated frequency of the allele “*a*” on the whole sample and *A* is the total number of alleles in the sample.

Population structure was assessed by DAPC (Discriminant Analysis of Principle Components) using posterior membership probabilities while assessing the membership stability by estimation of *a-scores*. Phenotypic and genotypic distance matrices were subjected to Mantel's correlation with permutation tests (Mantel, 1967). The minimum seed sample size required to capture 95% of alleles within an accession with a 95% certainty, during sampling for regeneration, was calculated as reported by Crossa (1989) for each accession. Considering the rarest biallelic locus (SNP), two alleles B_1 and B_2 with frequencies of p_1 and p_2 , so that ($p_1 + p_2 = 1$), the two possible outcomes will be,

$k_1 = B_1$ is not represented in the sample of *n* gametes

$k_2 = B_2$ is not represented in the sample of *n* gametes

Thus the probability of getting at least one copy of the each B_1 and B_2 will be $P(k_1^c \cap k_2^c)$,

$$P(k_1^c \cap k_2^c) = 1 - (1 - p_1)^n - (1 - p_2)^n$$

All the above-mentioned analyses were performed using R software v.3.6.0 (R Core Team, 2019). Custom scripted codes were used for filtering, distance matrix, heterozygosity estimations, and seed sample size computations, also packages from R-CRAN and GitHub like “ade4” (Jombart, 2008) and “ade4” (Dray et al., 2007) were used for computation of AMOVA and Mantel's test, respectively.

RESULTS

Phenotyping

Descriptive Statistics and *post-hoc* Tests

The variations in the mean and range estimates indicated considerable variability among landraces and wild accessions of sorghum, pearl millet and pigeonpea. The SNK test indicated significant ($p \leq 0.05$) mean differences among accessions (Supplementary Table 5). Levene's test indicated heterogeneous variances for all the quantitative traits in sorghum, pearl millet, and pigeonpea (Supplementary Table 6).

Phenotypic Diversity: Within and Between Accessions

The Gower's phenotypic distance matrix (Gower, 1971) was computed to obtain pairwise distances between plants of all the accessions. Within accession distances varied from 0.038 to 0.141,

0.145 to 0.271, and 0.071 to 0.410 for sorghum, pearl millet, and pigeonpea, respectively (Table 1). In sorghum IS 13215 (0.141) had the maximum mean within accession distance followed by IS 31637 (0.136) and IS 27325 (0.136), and the accession IS 12919 (0.038) showed the lowest within accession distance followed

by IS 13065 (0.046) and IS 2134 (0.048). In pearl millet, IP 12138 (0.271) showed the maximum within accession distance followed by the accessions IP 13112 (0.270) and IP 8761 (0.268), whereas the accession IP 21640 (0.145) had the lowest within accession distance followed by IP 22039 (0.159) and IP 21752

TABLE 1 | Mean phenotypic distances within and between accessions of sorghum, pigeonpea, and pearl millet.

Sorghum			Pigeonpea			Pearl millet		
Accession number	Within accession distances	Distance from other accessions	Accession number	Within accession distances	Distance from other accessions	Accession number	Within accession distances	Distance from other accessions
IS 12919	0.038	0.387	ICP 7035	0.071	0.351	IP 21640	0.145	0.302
IS 13065	0.046	0.318	ICP 11485	0.092	0.335	IP 22039	0.159	0.289
IS 2134	0.048	0.349	ICP 9124	0.094	0.314	IP 21752	0.162	0.317
IS 22407	0.050	0.335	ICP 9150	0.107	0.388	IP 6434	0.194	0.249
IS 22606	0.055	0.402	ICP 14059	0.120	0.316	IP 9446	0.205	0.259
IS 2348	0.060	0.364	ICP 13828	0.125	0.292	IP 11577	0.206	0.256
IS 14485	0.065	0.368	ICP 13628	0.132	0.252	IP 3616	0.207	0.247
IS 33844	0.069	0.342	ICP 11480	0.134	0.293	IP 13459	0.207	0.249
IS 32263	0.070	0.325	ICP 9877	0.143	0.285	IP 3389	0.207	0.245
IS 29605	0.073	0.318	ICP 7057	0.158	0.298	IP 5900	0.210	0.249
IS 13068	0.076	0.323	ICP 14296	0.179	0.263	IP 9824	0.212	0.269
IS 22428	0.076	0.346	ICP 13415	0.180	0.294	IP 10085	0.218	0.255
IS 35474	0.077	0.308	ICP 11491	0.181	0.275	IP 5441	0.218	0.257
IS 11005	0.077	0.365	ICP 9122	0.181	0.284	IP 11984	0.221	0.268
IS 18833	0.087	0.390	ICP 13575	0.189	0.269	IP 18147	0.223	0.269
IS 12965	0.088	0.415	ICP 13889	0.204	0.285	IP 17632	0.225	0.266
IS 34283	0.092	0.353	ICP 13316	0.205	0.305	IP 5253	0.228	0.293
IS 10897	0.092	0.350	ICP 6399	0.206	0.271	IP 6109	0.229	0.259
IS 14010	0.094	0.363	ICP 12190	0.212	0.277	IP 6244	0.234	0.258
IS 40238	0.096	0.347	ICP 11475	0.219	0.320	IP 4952	0.236	0.262
IS 18234	0.096	0.345	ICP 2309	0.227	0.314	IP 18157	0.236	0.260
IS 25476	0.097	0.354	ICP 14388	0.227	0.311	IP 19434	0.237	0.263
IS 3399	0.099	0.345	ICP 13546	0.228	0.284	IP 11677	0.238	0.274
IS 40031	0.099	0.357	ICP 12189	0.229	0.277	IP 20349	0.243	0.268
IS 35217	0.100	0.359	ICP 12041	0.235	0.291	IP 3269	0.243	0.279
IS 29508	0.102	0.346	ICP 16344	0.237	0.298	IP 12155	0.247	0.282
IS 2153	0.108	0.400	ICP 13999	0.242	0.293	IP 14071	0.249	0.264
IS 1128	0.109	0.378	ICP 14169	0.247	0.293	IP 7468	0.249	0.294
IS 40161	0.111	0.340	ICP 14233	0.263	0.294	IP 14418	0.250	0.278
IS 8330	0.112	0.414	ICP 15148	0.274	0.308	IP 6037	0.256	0.264
IS 21858	0.118	0.355	ICP 7621	0.274	0.292	IP 20407	0.257	0.301
IS 32252	0.118	0.370	ICP 10880	0.276	0.299	IP 13363	0.258	0.307
IS 13211	0.134	0.367	ICP 10889	0.299	0.317	IP 10705	0.265	0.276
IS 27325	0.136	0.338	ICP 12840	0.317	0.338	IP 8761	0.268	0.286
IS 31637	0.136	0.351	ICP 13545	0.410	0.384	IP 13112	0.270	0.281
IS 13215	0.141	0.395				IP 12138	0.271	0.275
Overall mean	0.090	0.387		0.203	0.302		0.227	0.271
Overall range	0.038–0.141	0.308–0.415		0.071–0.410	0.252–0.388		0.145–0.271	0.245–0.310

(0.162). In pigeonpea, ICP 13545 (0.410) showed the highest mean within accession distance followed by ICP 12840 (0.317) and ICP 10889 (0.299), whereas the least was noticed in ICP 7035 (0.071) followed by ICP 11485 (0.092) and ICP 9124 (0.094). The wild accessions of pearl millet had the minimum within accession distance [IP 21640 (0.145), IP 22039 (0.159), and IP 21752 (0.162)], in comparison to the overall scale of mean distance values of landraces (0.194–0.271). The same scenario was observed in sorghum where the wild accessions IS 14485 (0.065), IS 10897 (0.092), IS 11005 (0.077), IS 18833 (0.087), and IS 22428 (0.076) had low phenotypic within accession distances in comparison to the overall range of within accession distance values of landraces (0.038–0.141). On an average, distance among accessions was found to be higher than that of within accessions distance in all the three crops. Between accessions distance values were higher in accessions of sorghum (mean 0.387; range 0.308–0.415), followed by pigeonpea (mean 0.302; range 0.252–0.388), while low in pearl millet (mean 0.271; range 0.245–0.310) (**Table 1**).

Hierarchical clustering was constructed based on Gower's phenotypic distance, and the number of clusters was decided based on the number of accessions in each crop from which the data were collected. Thus, dendrogram trees were cut at 36 clusters for sorghum and pearl millet, and 35 clusters for pigeonpea, with the assumption that the individuals within accession clusters together. A cluster membership barplot was generated to visualize distribution or migration of individuals of different accessions to different clusters. The cluster wise stability was evaluated using the “clusterboot” function from the “fpc” package. The Jaccard coefficients between clusters of resampled data were >70 for 35 clusters in sorghum, 15 clusters in pigeonpea, and 16 clusters in pearl millet (**Supplementary Table 7**) and the remaining clusters showed values <70. The bootstrapped cluster dendrograms were plotted with approximately unbiased *p*-values (AU) and bootstrap probability (**Supplementary Figures 1A–C**) calculated using multiscale bootstrap resampling in the R-package “pvclust.” Bootstrap values were low in some cases of pearl millet and pigeonpea and this low bootstrap values would be a combined outcome of high variability in the data, large number of variable individuals, and the nature of clustering algorithm. Supporting the high variability and presence of valid clusters in the data, the “pvpick” function from the “pvclust” R-package yielded 66, 738, and 109 significant clusters in sorghum, pearl millet, and pigeonpea, respectively. Thus, the presence of large number of significant clusters within the studied accessions illustrates the higher variability for the observed traits and ultimately represents the higher diversity within the studied landraces. In sorghum, except cluster numbers 7, 13, and 14 all other 33 clusters have shown exclusive clustering of each accession into singleton clusters (**Figure 2A**). In cluster number 7, individuals of entries IS 8330 and IS 12965 were found to clustered together. The individuals of accession IS 2153 were found to be distributed in two clusters (58 individuals in cluster 14 and 139 individuals in cluster 13). In pearl millet and pigeonpea, clustering patterns showed that in most accessions, the individuals were not clustered uniquely, and

found mixed with other accessions. In pigeonpea all the individuals of three accessions ICP 9150, ICP 7035, and ICP 11485 were clustered in clusters 1, 10, and 14, respectively. However, in ICP 9124, except a single individual all the other individuals were clustered in cluster 16 (**Figure 2B**). In pearl millet, no exclusive clusters were observed and all the 36 clusters showed mixing of individuals from different accessions (**Figure 2C**). However, The majority of individuals of wild accessions (IP 21640, IP 21752, and IP 22039) were distributed in 3 clusters (C-1, C-2, and C-3) showing their phenotypic similarity.

Genotypic Diversity

After filtering, we obtained 45,249 SNPs from a total of 76,753 SNPs in pearl millet, 19,052 SNPs from a total of 38,898 SNPs in sorghum, and 8,211 SNPs from a total of 10,096 SNPs in pigeonpea. The SNPs displayed good coverage across genome in all the three crops (**Figure 3**). Over the 10 chromosomes of sorghum the number of SNPs ranged from 909 to 2,988, and over the 7 chromosomes of pearl millet the number of SNPs ranged from 5,086 to 6,639, and from 121 to 755 over the 11 chromosomes of pigeonpea. The information of number of SNPs in each chromosome of sorghum, pearl millet and pigeonpea is presented in **Supplementary Table 8**.

AMOVA

The analysis of molecular variance (Excoffier et al., 1992) was performed by providing predefined populations, that each accession as a separate population. The results showed that the proportion of molecular variance contributed by within accession variance depicted a low value of 26.3% in sorghum, a relatively higher value in pigeonpea (57.0%), and the highest in pearl millet (80.2%) (**Table 2**; **Figure 4**). Variance among populations was high in sorghum (73.7%), while low in pearl millet (19.8%) and intermediate in pigeonpea (43%).

Genotypic Diversity: Within and Between Accessions

Modified Rogers Distance (MRD) (Wright, 1978; Goodman and Stuber, 1983) between pairs of individuals were estimated. Pairwise MRD within each accession was averaged, thus the overall mean genetic distance within each accession varied from 0.031 (IS 33844) to 0.342 (IS 18833), 0.181 (IP 9824) to 0.300 (IP 22039), and 0.040 (ICP 9150) to 0.393 (ICP 10889) in sorghum, pearl millet, and pigeonpea, respectively (**Table 3**). Three of the five wild accessions studied in sorghum namely IS 18833 (0.342), IS 14485 (0.329), and IS 10897 (0.316), showed higher within accession distance values relative to the studied landraces and the other two wild accessions, IS 11005 (0.119) and IS 22428 (0.127), showed midrange values. However, all the studied wild accessions of pearl millet, IP 21752 (0.273), IP 21640 (0.279), and IP 22039 (0.300), showed higher within accession distance values relative to the mean distances of the pearl millet landraces studied. Averaging the MRD among accessions were found to be higher in comparison to within accession distances. Comparing the three crops, higher scale

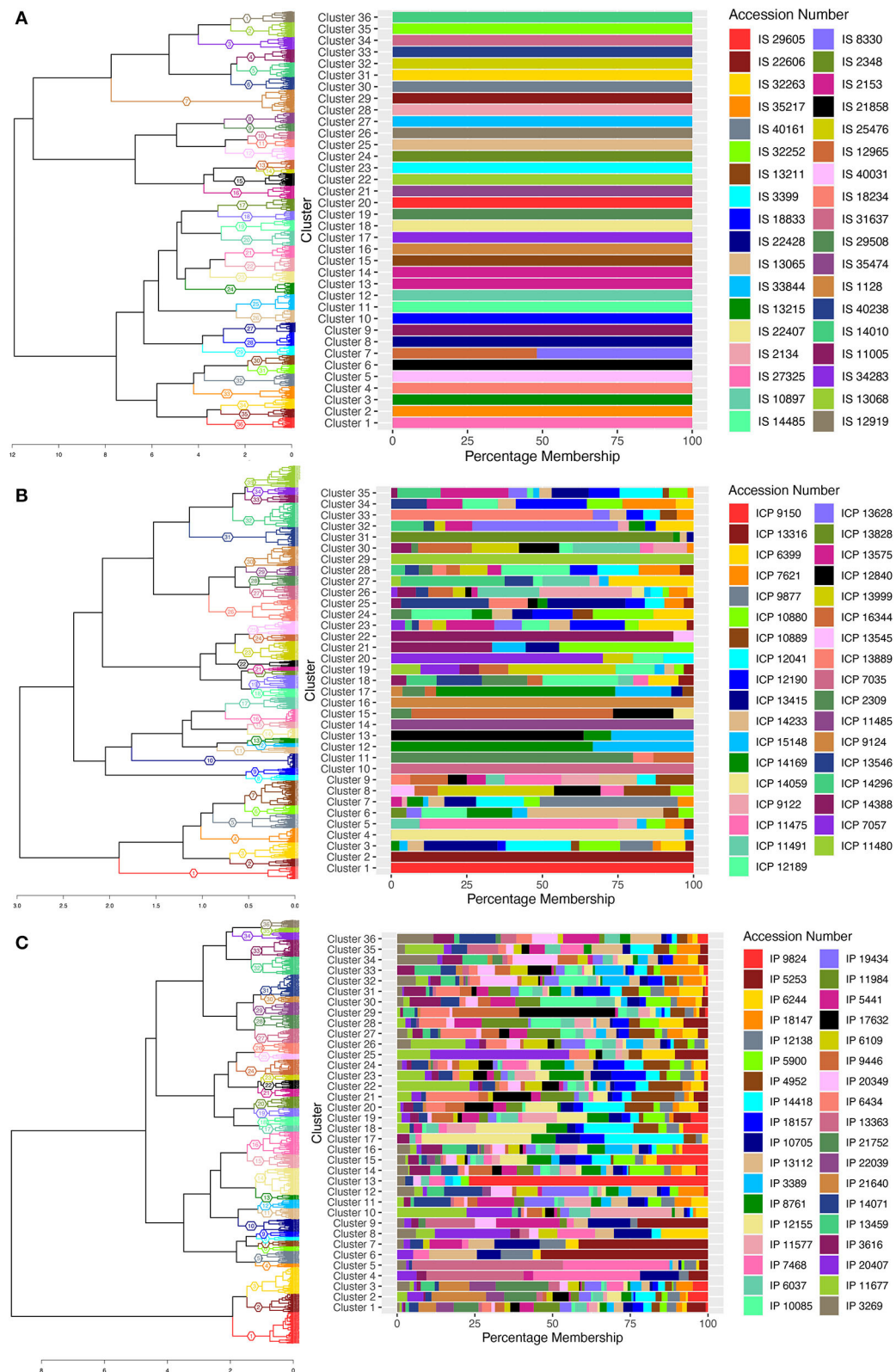


FIGURE 2 | Cluster dendrogram of single plant phenotypic distances, using ward.D2 clustering algorithm, 36, 35, and 36 clusters for sorghum, pigeonpea, and pearl millet, respectively, represented with colors and cluster numbers, with percentage membership of accessions into each cluster denoted by colors in the adjacent bar graph. **(A)** the cluster dendrogram for sorghum, **(B)** the cluster dendrogram for pigeonpea, and **(C)** the cluster dendrogram for pearl millet.

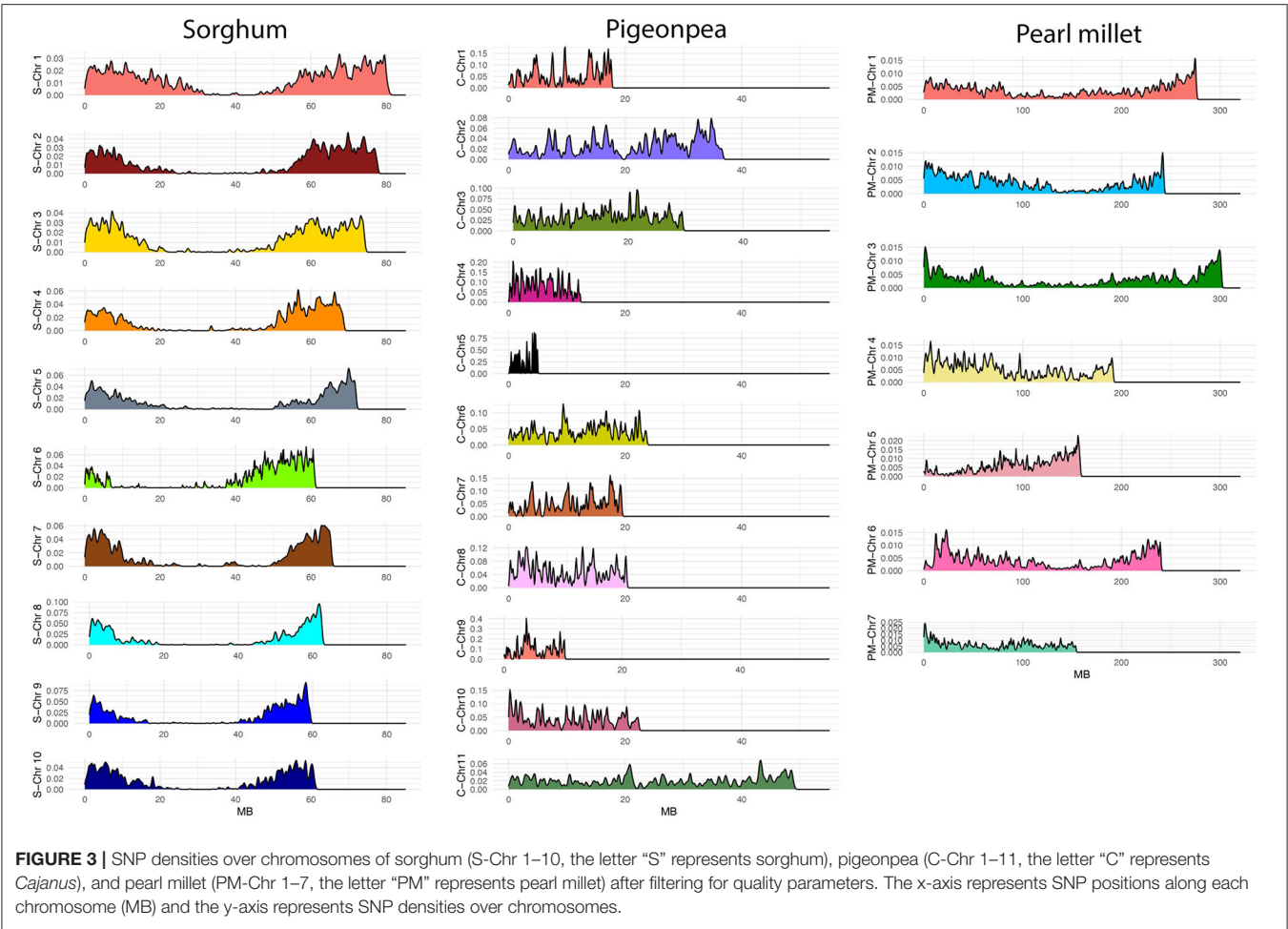


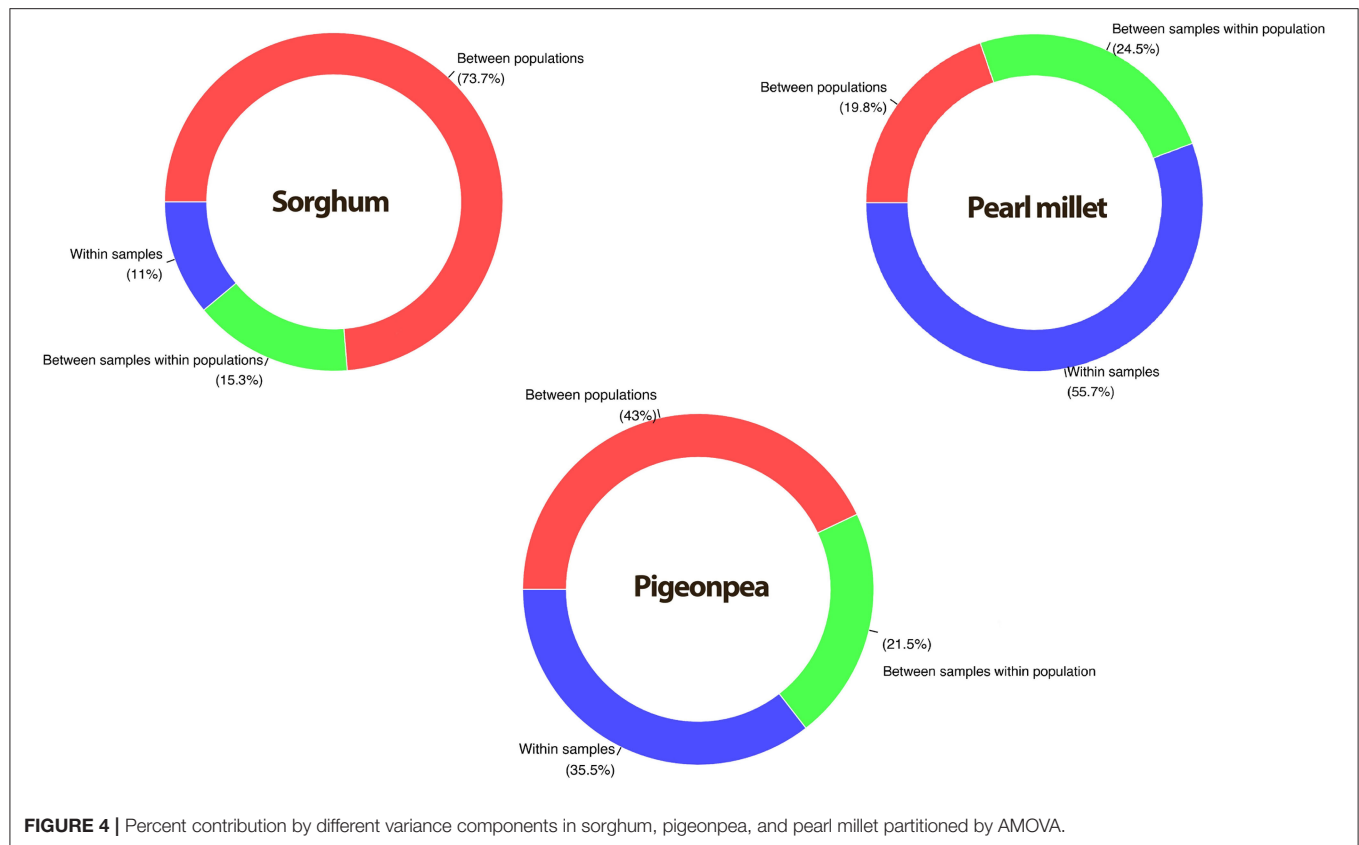
FIGURE 3 | SNP densities over chromosomes of sorghum (S-Chr 1–10, the letter “S” represents sorghum), pigeonpea (C-Chr 1–11, the letter “C” represents *Cajanus*), and pearl millet (PM-Chr 1–7, the letter “PM” represents pearl millet) after filtering for quality parameters. The x-axis represents SNP positions along each chromosome (MB) and the y-axis represents SNP densities over chromosomes.

TABLE 2 | AMOVA on DArTSeq- SNP data of sorghum, pigeonpea, and pearl millet assuming each accession as a single population.

Variance components	Df	Sum Sq	Mean Sq	Variance %	Sigma	Phi	P-value
Sorghum							
Between populations	35	2,508,925	71683.5	73.7	2336.9	0.8895	0.01
Between samples within populations	506	665,752	1315.7	15.3	482.8	0.5797	0.01
Within samples	542	189,717	350.0	11.0	350.0	0.7373	0.01
Total	1083	3,364,396	3106.5	100	3169.8		
Pigeonpea							
Between populations	35	368,370	10524.8	43.0	331.2	0.6449	0.01
Between samples within populations	503	304,494	605.3	21.5	165.8	0.3773	0.01
Within samples	539	147,500	273.6	35.5	273.6	0.4297	0.01
Total	1077	820,366	761.7	100	770.7		
Pearl millet							
Between populations	35	2,294,552	65558.6	19.8	1061	0.4427	0.01
Between samples within populations	981	5,507,738	5614.4	24.5	1313.1	0.3052	0.01
Within samples	1017	3,038,922	2988.1	55.7	2988.1	0.1978	0.01
Total	2033	10,841,212	5332.6	100	5362.3		

of between accession distance values were found in sorghum (0.360–0.435), followed by pigeonpea (0.237–0.422) and pearl millet (0.276–0.324).

Heterozygosity among accessions in sorghum varied from 0.019 in IS 31637 to 0.159 in IS 18833. Among five wild accessions studied in sorghum, three accessions had higher heterozygosity



(0.111 in IS 10897, 0.120 in IS 14485, and 0.159 in IS 18833) in comparison to all the landraces, while the other two wild accessions had low heterozygosity (0.026 in IS 22428 and 0.051 in IS 11005). In pearl millet, IP 9824 (0.051) and IP 22039 (0.137) estimated the lowest and highest heterozygosity, respectively. As in sorghum, the wild accessions of pearl millet *viz.*, IP 21640 (0.134), IP 21752 (0.118), and IP 22309 (0.137) depicted maximal heterozygosity estimates in comparison to the landraces. In pigeonpea, ICP 9150 (0.007) had the minimum heterozygosity while ICP 10889 (0.187) had the highest heterozygosity (Table 3).

Clustering based on MRD grouped the individuals of all the accessions into different clusters. The cluster wise stability was assessed using the “clusterboot” function from R-package “fpc” with 100 bootstraps. About 33 clusters in sorghum, 24 clusters in pigeonpea and 19 clusters in pearl millet showed Jaccard coefficient values >70 and all the other clusters showed values <70. The Jaccard coefficient values of all the clusters were presented in **Supplementary Table 9**. Bootstrapped dendrograms (**Supplementary Figures 2A–C**) with 100 bootstraps were plotted using the “aboot” function provided in the R-package “poppr.” The dendrogram tree was cut at 36 clusters considering number of accessions in the respective crops, with an assumption that the individuals of each accession should aggregate into singleton cluster. Also a cluster membership bar-plot was used to visualize this cluster partition and migration of plants to different clusters. In sorghum, 19 accessions, *viz.*, IS 1128, IS 12919, IS 12965, IS 13065, IS 18234,

IS 18833, IS 2153, IS 21858, IS 22428, IS 25476, IS 31637, IS 32252, IS 32263, IS 33844, IS 3399, IS 34283, IS 13068, IS 35474, and IS 40031 were found to be uniform, by clustering of all the individual of an accessions into separate singleton clusters (**Figure 5A**), while the landraces IS 29508 and IS 29605 were found to be grouped in a single cluster. All other accessions of sorghum were found to have mixtures. In pigeonpea, ICP 2309, ICP 9124, ICP 7057, ICP 9877, ICP 11480, ICP 14059, ICP 13628, ICP 7035, ICP 9150, ICP 13828, and ICP 15122 showed perfect singleton clustering (**Figure 5B**), while other accessions showed overlapping of individuals of different accessions which may be explained due to the heterogeneity achieved in evolutionary gene-flow or the presence of admixtures in the respective accessions. In pearl millet, a completely distinctive and complex distribution of accessions into clusters has been noticed. The accessions IP 9824, IP 7468, IP 11577, IP 11677, IP 13363, and IP 19434 showed perfect singleton clustering (**Figure 5C**) while all other accessions were not clustered uniquely to singleton clusters, indicating heterogeneity within landraces and sharing of alleles between accessions. The wild accessions of pearl millet showed an interesting pattern of clustering that the individuals of the accession IP 22039 was shared between cluster numbers 11 and 12 showing the presence of a variable set of alleles or two subpopulations, and also there can be seen some individuals of the accession IP 21752 clustered with the individuals of the accession IP 21640 in cluster number 14 depicting some similar alleles between these two accessions.

TABLE 3 | Mean genotypic within and between accession distances, observed heterozygosity within accessions (heWs) and estimated seed sample sizes using the least DArTSeq–SNP allelic frequency in all the accessions of sorghum, pigeonpea, and pearl millet.

Sorghum					Pigeonpea					Pearl millet				
Accession number	Within accession distance	Distance from other accessions	Heterozygosity	No. of seeds	Accession Number	Within accession distance	Distance from other accessions	Heterozygosity	No. of seeds	Accession number	Within accession distance	Distance from other accessions	Heterozygosity	No. of seeds
IS 10897	0.316	0.389	0.111	89	ICP 10880	0.348	0.334	0.168	89	IP 10085	0.255	0.292	0.092	173
IS 11005	0.119	0.405	0.051	101	ICP 10889	0.393	0.422	0.187	77	IP 10471	0.254	0.285	0.088	173
IS 1128	0.130	0.393	0.037	89	ICP 11475	0.201	0.283	0.058	89	IP 10705	0.265	0.286	0.104	161
IS 12919	0.038	0.381	0.032	89	ICP 11480	0.196	0.276	0.051	89	IP 11577	0.256	0.288	0.092	167
IS 12965	0.035	0.384	0.032	89	ICP 11485	0.066	0.283	0.017	89	IP 11677	0.267	0.288	0.106	161
IS 13065	0.094	0.365	0.033	89	ICP 11491	0.197	0.286	0.052	89	IP 11984	0.241	0.277	0.092	167
IS 13068	0.037	0.367	0.031	89	ICP 12041	0.169	0.242	0.043	89	IP 12138	0.270	0.285	0.103	167
IS 13211	0.145	0.400	0.047	89	ICP 12189	0.230	0.254	0.071	89	IP 12155	0.253	0.280	0.098	161
IS 13215	0.244	0.387	0.084	77	ICP 12190	0.186	0.244	0.053	89	IP 13112	0.225	0.280	0.085	179
IS 14010	0.216	0.367	0.076	83	ICP 12840	0.232	0.254	0.072	89	IP 13363	0.202	0.283	0.064	161
IS 14485	0.329	0.396	0.120	89	ICP 13316	0.228	0.272	0.123	89	IP 13459	0.225	0.285	0.085	161
IS 18234	0.034	0.379	0.024	89	ICP 13415	0.096	0.244	0.016	89	IP 14418	0.243	0.276	0.100	185
IS 18833	0.342	0.435	0.159	89	ICP 13545	0.145	0.246	0.034	89	IP 17632	0.218	0.276	0.085	167
IS 2134	0.240	0.373	0.090	89	ICP 13546	0.181	0.244	0.054	89	IP 18147	0.231	0.286	0.083	155
IS 2153	0.113	0.379	0.048	89	ICP 13575	0.185	0.245	0.053	89	IP 18157	0.251	0.291	0.088	161
IS 21858	0.104	0.376	0.044	89	ICP 13628	0.153	0.256	0.033	89	IP 19434	0.237	0.277	0.088	167
IS 22407	0.165	0.378	0.070	89	ICP 13828	0.040	0.245	0.009	89	IP 20349	0.246	0.284	0.101	197
IS 22428	0.127	0.387	0.026	89	ICP 13889	0.105	0.247	0.025	89	IP 20407	0.246	0.281	0.095	161
IS 22606	0.131	0.380	0.064	77	ICP 13999	0.210	0.253	0.058	89	IP 21640	0.279	0.310	0.134	161
IS 2348	0.085	0.390	0.043	89	ICP 14059	0.085	0.250	0.014	83	IP 21752	0.273	0.311	0.118	167
IS 25476	0.114	0.395	0.034	89	ICP 14169	0.163	0.246	0.051	89	IP 22039	0.300	0.324	0.137	161
IS 27325	0.310	0.381	0.102	47	ICP 14233	0.194	0.246	0.059	89	IP 3269	0.230	0.282	0.090	161
IS 29508	0.128	0.360	0.042	89	ICP 14296	0.166	0.246	0.040	89	IP 3389	0.249	0.282	0.098	179
IS 29605	0.099	0.361	0.036	89	ICP 14388	0.132	0.255	0.039	89	IP 3616	0.259	0.280	0.106	185
IS 31637	0.036	0.400	0.019	89	ICP 15122	0.050	0.237	0.009	89	IP 4952	0.264	0.281	0.108	173
IS 32252	0.245	0.396	0.096	89	ICP 15148	0.199	0.243	0.062	89	IP 5253	0.238	0.286	0.081	155
IS 32263	0.035	0.395	0.029	89	ICP 16344	0.204	0.268	0.063	89	IP 5441	0.248	0.280	0.101	167
IS 33844	0.031	0.381	0.024	89	ICP 2309	0.202	0.274	0.064	89	IP 5900	0.254	0.288	0.105	155
IS 3399	0.038	0.362	0.030	89	ICP 6399	0.220	0.271	0.067	89	IP 6037	0.242	0.281	0.092	173
IS 34283	0.048	0.377	0.034	89	ICP 7035	0.043	0.285	0.011	89	IP 6109	0.246	0.280	0.103	203
IS 35217	0.130	0.373	0.040	89	ICP 7057	0.143	0.265	0.043	89	IP 6244	0.262	0.286	0.102	167
IS 35474	0.109	0.364	0.040	89	ICP 7621	0.243	0.265	0.090	89	IP 6434	0.246	0.282	0.096	167

(Continued)

TABLE 3 | Continued

Accession number	Sorghum				Pigeonpea				Pearl millet			
	Within accession distance	Distance from other accessions	Heterozygosity	No. of seeds	Within accession distance	Distance from other accessions	Heterozygosity	No. of seeds	Within accession distance	Distance from other accessions	Heterozygosity	No. of seeds
IS 40031	0.040	0.389	0.038	89	0.187	0.280	0.054	89	0.217	0.276	0.088	167
IS 40161	0.105	0.384	0.033	89	0.053	0.243	0.010	89	0.237	0.279	0.088	155
IS 40238	0.204	0.372	0.065	89	0.040	0.241	0.007	89	0.226	0.278	0.084	179
IS 8330	0.129	0.382	0.056	89	0.148	0.251	0.031	89	0.181	0.288	0.051	155
Overall mean	0.135	0.383	0.054	87	0.168	0.264	0.053	88	0.245	0.285	0.095	168
Overall range	0.031–0.342	0.360–0.435	0.019–0.159	47–101	0.04–0.393	0.237–0.422	0.007–0.187	77–89	0.181–0.3	0.276–0.324	0.051–0.137	155–203

Shannon Diversity

The Shannon diversity was estimated for all 36 accessions in sorghum, pigeonpea, and pearl millet (Table 4). The values ranged from 0.113 to 2.363 with an overall mean of 0.275 in sorghum, 0.121–1.739 with an overall mean of 0.498 in pigeonpea and 1.128–2.715 with an overall mean of 1.856 in pearl millet. In sorghum, the accession IS 22606 (0.113) had the lowest Shannon diversity followed by the accessions IS 35474 (0.114) and IS 35217 (0.116), and the accession IS 18833 (2.363) had the highest Shannon diversity followed by accessions IS 32252 (1.138) and IS 14485 (0.583). In pigeonpea the accession ICP 9150 (0.121) had the lowest value of Shannon diversity followed by the accessions ICP 15122 (0.148) and ICP 13415 (0.156), and the accession ICP 13316 (1.739) had the highest Shannon diversity followed by the accessions ICP 10880 (1.542) and ICP 10889 (1.404). In pearl millet the accessions IP 9824 (1.128), IP 10471 (1.270), and IP 18157 (1.335) had lower values of Shannon diversity and the accessions IP 21640 (2.715), IP 6109 (2.677), and IP 20349 (2.489) had higher values. It should be noted that, in both sorghum and pearl millet, all the wild accessions had high values of Shannon diversity. Three out of five wild accessions in sorghum IS 18833 (2.363), IS 14485 (0.583), and IS 10897 (0.495) had higher values relative other accessions of sorghum, while the remaining two wild accessions IS 22428 (0.192) and IS 11005 (0.198) found to have intermediate values. Also, the wild accessions of pearl millet had relatively higher values of Shannon diversity IP 21640 (2.715), IP 22039 (2.382), and IP 21752 (2.359) in comparison to all other accessions of pearl millet.

Relationship Between Phenotypic and Genotypic Distances

Mantel's correlation between phenotypic and genotypic distance matrices showed highly significant positive correlation ($r = 0.45$, $P \leq 0.01$) for sorghum, pearl millet ($r = 0.13$, $P \leq 0.01$), and pigeonpea ($r = 0.19$, $P \leq 0.01$), thus depicting the effectiveness of complimentary use of molecular and phenotypic tools as a better approach for the assessment of the genetic diversity.

Population Structure Using DAPC

Detecting the number of clusters using the *find.cluster* function hasn't shown any significant elbow of reduction in BIC values (Supplementary Figure 3), instead, a gradual reduction in the BIC values was seen on increasing number of clusters. So that, DAPC was carried out using 36 clusters representing the 36 accessions in sorghum, pearl millet and pigeonpea. The *optm.a.score* function detected an optimal first 45 PCs for pearl millet and first 7 PCs for both sorghum and pigeonpea. Based on the posterior membership probabilities the population membership graph showing the population structure was created. In sorghum (Supplementary Figure 4A), 17 accessions (IS 1128, IS 12965, IS 18234, IS 18833, IS 2153, IS 21858, IS 25476, IS 31637, IS 32252, IS 32263, IS 33844, IS 3399, IS 34283, IS 35217, IS 35474, IS 40031, and IS 40161) were clustered exclusively into separate populations, whereas both the accessions IS 29508 and IS 29605 were clustered into a single population and all the other accessions are seen to have mixtures at different levels. In the

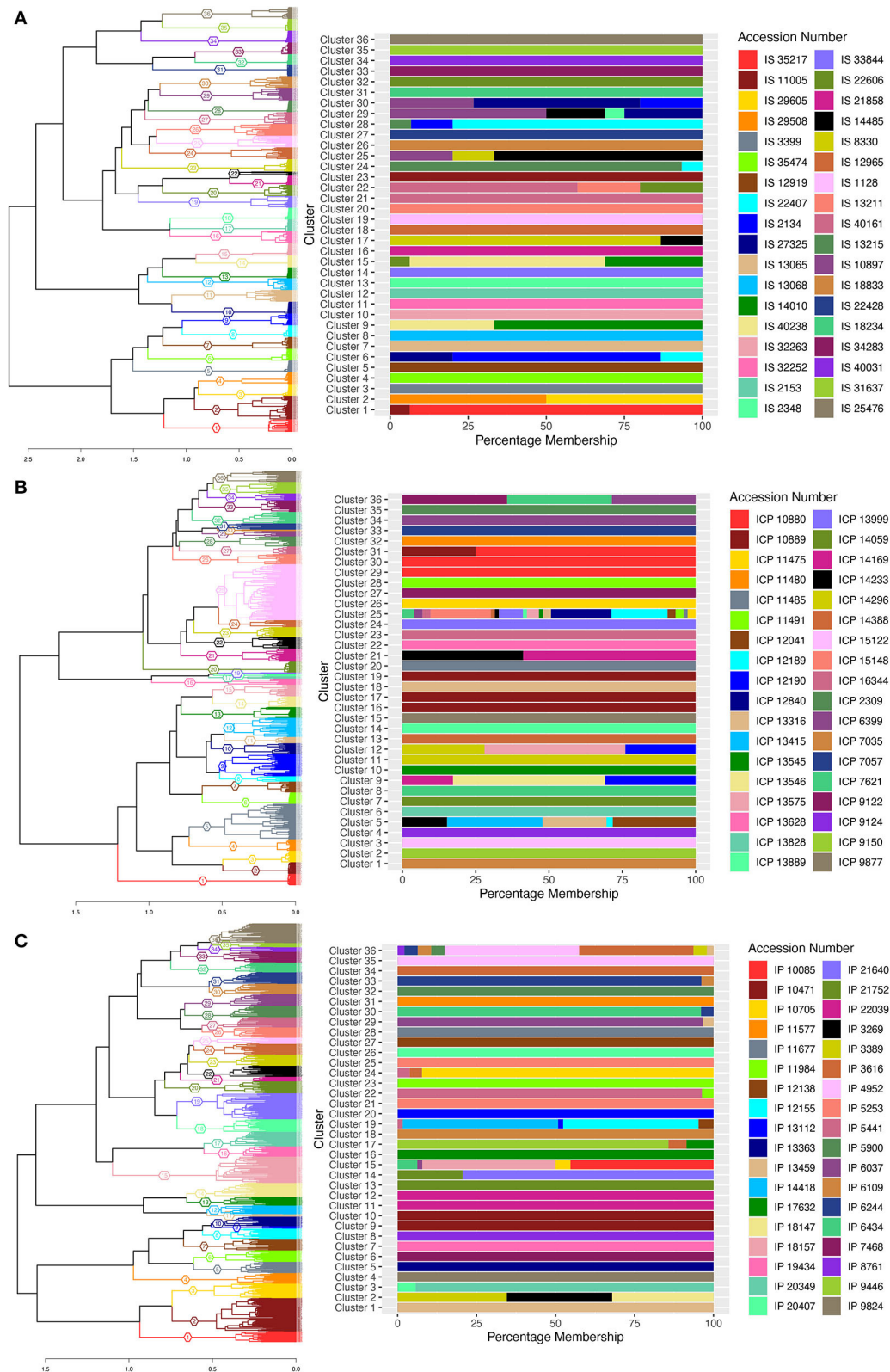


FIGURE 5 | Cluster dendrogram of single plant genotypic distances estimated from DARTSeq- SNP data, using ward.D2 clustering algorithm, 36 clusters for sorghum, pigeonpea, and pearl millet, represented with colors and cluster numbers, with percentage membership of accessions into each cluster denoted by colors in the adjacent bar graph. **(A)** the cluster dendrogram of sorghum, **(B)** the cluster dendrogram of pigeonpea, and **(C)** the cluster dendrogram of pearl millet.

TABLE 4 | Shannon diversity (H') estimates within each accession of sorghum, pigeonpea, and pearl millet estimated from the DArTSeq-SNPs.

Sorghum		Pigeonpea		Pearl millet	
Accession number	H'	Accession number	H'	Accession number	H'
IS 22606	0.113	ICP 9150	0.121	IP 9824	1.128
IS 35474	0.114	ICP 15122	0.148	IP 10471	1.270
IS 35217	0.116	ICP 13415	0.156	IP 18157	1.335
IS 13068	0.117	ICP 13828	0.161	IP 13363	1.357
IS 14010	0.117	ICP 9124	0.167	IP 11577	1.408
IS 18234	0.119	ICP 14059	0.168	IP 5253	1.430
IS 3399	0.120	ICP 7035	0.212	IP 12138	1.458
IS 33844	0.120	ICP 9877	0.239	IP 10085	1.518
IS 12965	0.120	ICP 13628	0.261	IP 8761	1.531
IS 12919	0.123	ICP 14296	0.282	IP 18147	1.664
IS 34283	0.128	ICP 11485	0.312	IP 13459	1.685
IS 31637	0.129	ICP 11491	0.313	IP 10705	1.706
IS 13065	0.131	ICP 13889	0.318	IP 12155	1.706
IS 40238	0.132	ICP 13545	0.321	IP 11677	1.709
IS 22407	0.132	ICP 9122	0.342	IP 6037	1.710
IS 32263	0.134	ICP 13999	0.397	IP 6244	1.758
IS 25476	0.135	ICP 14388	0.398	IP 20407	1.797
IS 40031	0.136	ICP 12041	0.404	IP 19434	1.802
IS 40161	0.138	ICP 11480	0.419	IP 11984	1.809
IS 1128	0.147	ICP 13575	0.449	IP 9446	1.925
IS 2348	0.155	ICP 6399	0.449	IP 5900	1.952
IS 29605	0.159	ICP 14233	0.473	IP 6434	1.972
IS 13211	0.167	ICP 12190	0.532	IP 7468	1.975
IS 29508	0.179	ICP 11475	0.541	IP 17632	1.978
IS 8330	0.185	ICP 12840	0.544	IP 5441	2.019
IS 2134	0.189	ICP 15148	0.561	IP 3389	2.026
IS 22428	0.192	ICP 12189	0.575	IP 3269	2.059
IS 11005	0.198	ICP 13546	0.587	IP 4952	2.076
IS 2153	0.234	ICP 16344	0.623	IP 3616	2.133
IS 27325	0.288	ICP 7057	0.626	IP 13112	2.256
IS 21858	0.390	ICP 2309	0.667	IP 21752	2.359
IS 13215	0.494	ICP 14169	0.679	IP 14418	2.379
IS 10897	0.495	ICP 7621	0.798	IP 22039	2.382
IS 14485	0.583	ICP 10889	1.404	IP 20349	2.498
IS 32252	1.138	ICP 10880	1.542	IP 6109	2.677
IS 18833	2.363	ICP 13316	1.739	IP 21640	2.715
Overall mean	0.275		0.498		1.865
Overall range	0.113–2.363		0.121–1.739		1.128–2.715

population structure of pigeonpea (**Supplementary Figure 4B**), seven accessions (ICP 13828, ICP 14059, ICP 7035, ICP 11485, ICP 14169, ICP 15122, and ICP 9150) were found to be pure, and in pearl millet (**Supplementary Figure 4C**), seven accessions (IP 6434, IP 7468, IP 9824, IP 18157, IP 13363, and IP 3389) were

seen to be clustered perfectly without any posterior probability for assessment to other populations while all the other accessions in both pearl millet and pigeonpea have a considerable amount of mixtures depicting the heterogeneity in the respective accessions.

For all the 36 populations, the quality of the attribution of accessions into populations were investigated by estimating the *a*-scores (**Supplementary Table 10**). An *a*-score of 1 represents an accurate allocation of the plants into groups. Twenty-seven clusters (1, 3, 4, 5, 6, 7, 8, 10, 11, 13, 14, 15, 17, 18, 19, 21, 22, 23, 24, 27, 28, 29, 30, 33, 34, 35, and 36) in sorghum, 3 clusters in pearl millet (4, 24, and 30) and 17 clusters in pigeonpea (1, 2, 10, 11, 12, 13, 14, 15, 16, 17, 21, 22, 24, 25, 29, 32, and 33) showed a greater reliability (*a*-score = 0.81–0.99). The clusters (2, 9, 12, 16, 26, and 31) in sorghum, the clusters (2, 5, 6, 8, 10, 12, 13, 15, 16, 17, 18, 19, 20, 21, 25, 26, 27, 28, and 29) in pearl millet and the clusters (4, 5, 6, 7, 9, 18, 19, 20, 23, 26, 30, 31, 34, and 36) in pigeonpea showed an average reliability (*a*-score = 0.65–0.80) in the attribution, whereas 3 clusters (20, 25, and 32) from sorghum, 14 clusters (1, 3, 7, 9, 11, 14, 22, 23, 31, 32, 33, 34, 35, and 36) from pearl millet and 5 clusters (3, 8, 27, 28, and 35) from pigeonpea were found to have a low reliability (*a*-score = < 0.65) in the attribution to the DAPC detected populations.

Estimation of Seed Sample Size

Seed sample sizes required for regeneration to capture 95% of the alleles with an expected probability of 95%, was estimated based on the allelic frequencies of the DArTSeq-SNPs, for each accession using the model proposed by Crossa (1989). The results of the sample sizes required are given in **Table 3**. Seed sample sizes for sorghum ranged from 47 to 101, 155 to 203 for pearl millet, and 77 to 89 for pigeonpea. The seed sample size increments exponentially after the alternate allele frequency attains a value below 0.1 (**Figure 6**) depicting the need for an exponentially larger sample size for conserving the alleles with frequencies below 0.1. The number of rare allelic variants or markers (frequency less than or equal to 5% within accessions) preserved in the recommended sample size for each accession of the three crops (**Supplementary Table 11**) ranged from 345 to 3,075 in sorghum, 231 to 878 in pigeonpea, and 3,444 to 6,726 in pearl millet.

DISCUSSION

Sorghum, pearl millet, and pigeonpea are the important food crops, providing food and income to a large population thriving in the arid and semi-arid tropics. However, in this era of modern agriculture, landraces of these crops are becoming prone to genetic erosion, (Hammer et al., 1996; Shewayrga et al., 2008; Pattanashetti et al., 2016). Most landraces, that were permanently extinct from the farmers' field over the course of agricultural development, are only available in genebanks' collections. As each landrace possess a unique genetic fingerprint of ages of acclimatization to diverse environmental conditions, they are considered as an indispensable source of genetic variations by plant breeders and can address a potential scope in the development of improved varieties with higher productivity, nutrients, and climate resilience, etc. (Dwivedi et al., 2016).

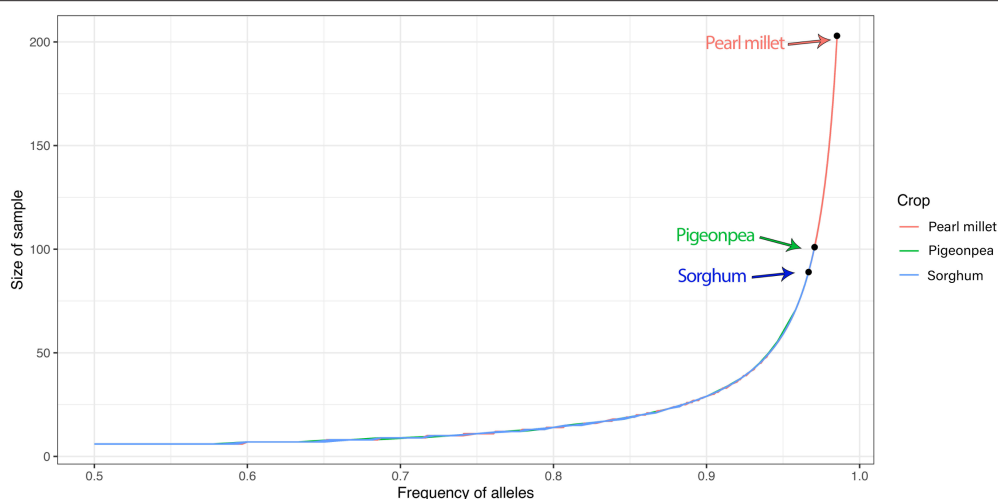


FIGURE 6 | Plot of seed sample sizes needed for regeneration vs. allelic frequencies to preserve 95% of alleles showing the exponential increase in the size of sample required with the decrease in the alternate allele frequency. The arrows represent the highest values of seed sample size estimated in sorghum, pigeonpea, and pearl millet.

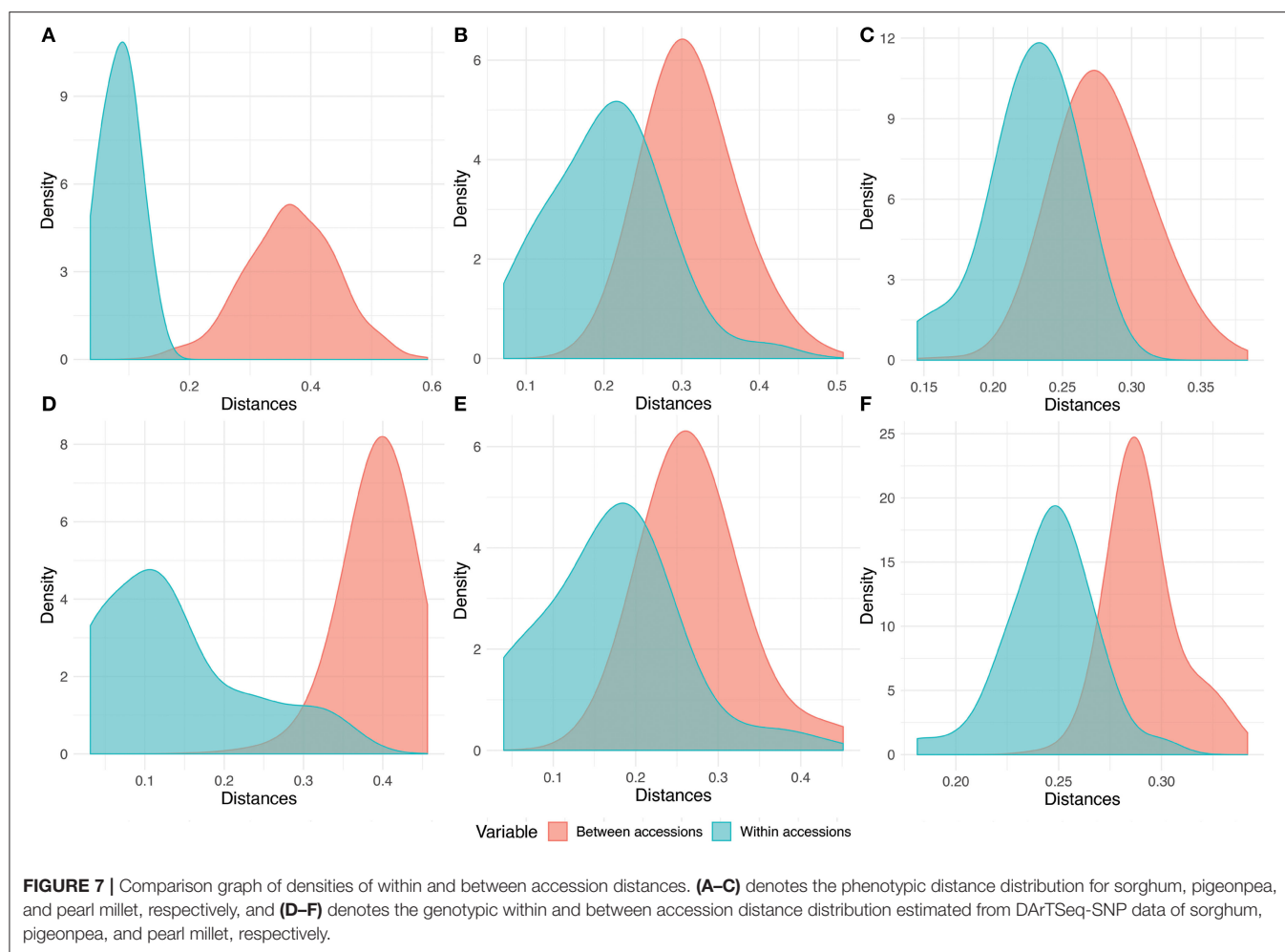
Thus, conserving landraces with their inherent genetic variability is crucial for ensuring food security in the near future and also for sustainable agriculture. ICRISAT genebank conserves about 42,000 accessions of sorghum, 24,000 accessions of pearl millet, and over 13,000 accessions of pigeonpea, wherein about 86% of sorghum and pearl millet collections and over 60% of pigeonpea collections are landraces. The main focus of genebank curator is to maintain the genetic integrity and diversity within accessions while regeneration. Hence, this study assessed the diversity within landrace accessions by phenotyping and genotyping a large number of plants within each accession and estimated the seed sample size required in order to conserve the inherent diversity.

Enormous variability was observed within and among landraces of sorghum, pigeonpea, and pearl millet. Molecular variance within accessions was observed to be low in sorghum (26.3%), highest in pearl millet (80.2%), while pigeonpea showing an intermediate within accession variance of 57.0%. Our results are in correspondence with previous works, by various authors (Tostain et al., 1987; Tostain and Marchais, 1989; Busso et al., 2000; Bashir et al., 2015) on pearl millet landraces, reported a high intra-population variation of 70–90% and higher observed heterozygosity of 0.77–0.82. However, Bhattacharjee et al. (2002) reported a low 30.89% within accession variability using RFLP markers in pearl millet, also the author addressed this low variability as a contradiction for a cross-pollinated crop like pearl millet and discussed various instances that would have caused this lower variability. In sorghum, Adugna (2014) reported a 54.44% molecular variance due to diversity within landrace populations that were conserved on farms in Ethiopia. No studies investigating landraces diversity within accessions were reported in sorghum and pigeonpea, while few studies are on landraces conserved on-farm that are continually evolving through outcrossing and selections (Djè et al., 1999; D'Andrea

and Casey, 2002; Songok et al., 2010; Adugna, 2014; Bashir et al., 2015).

The phenotypic and genotypic within accession distances were scaled toward the higher values in pearl millet, so that blurring the differentiation of within and between accessions diversity. The density distribution of within and between accession distances in pearl millet showed this scenario clearly, exhibiting the merging of densities (Figures 7C,F) of within and between accession distances in both phenotypic and genotypic evaluation. Pigeonpea being often cross-pollinated also depicted a pattern of overlapping within and among accession distances in both phenotypic and genotypic evaluation (Figures 7B,E). Whereas, sorghum showed a clear separation of distances within accessions from distances between accessions in both phenotypic and genotypic assessment, depicting the higher uniformity and homogeneity within the accessions (Figures 7A,D). The higher values and merging of between and within accession distances in pearl millet and pigeonpea shows the high phenotypic and genotypic heterogeneity within accessions, and also the clear separation of densities of within and between accession distances in sorghum clearly explains the higher uniformity within the accessions of sorghum.

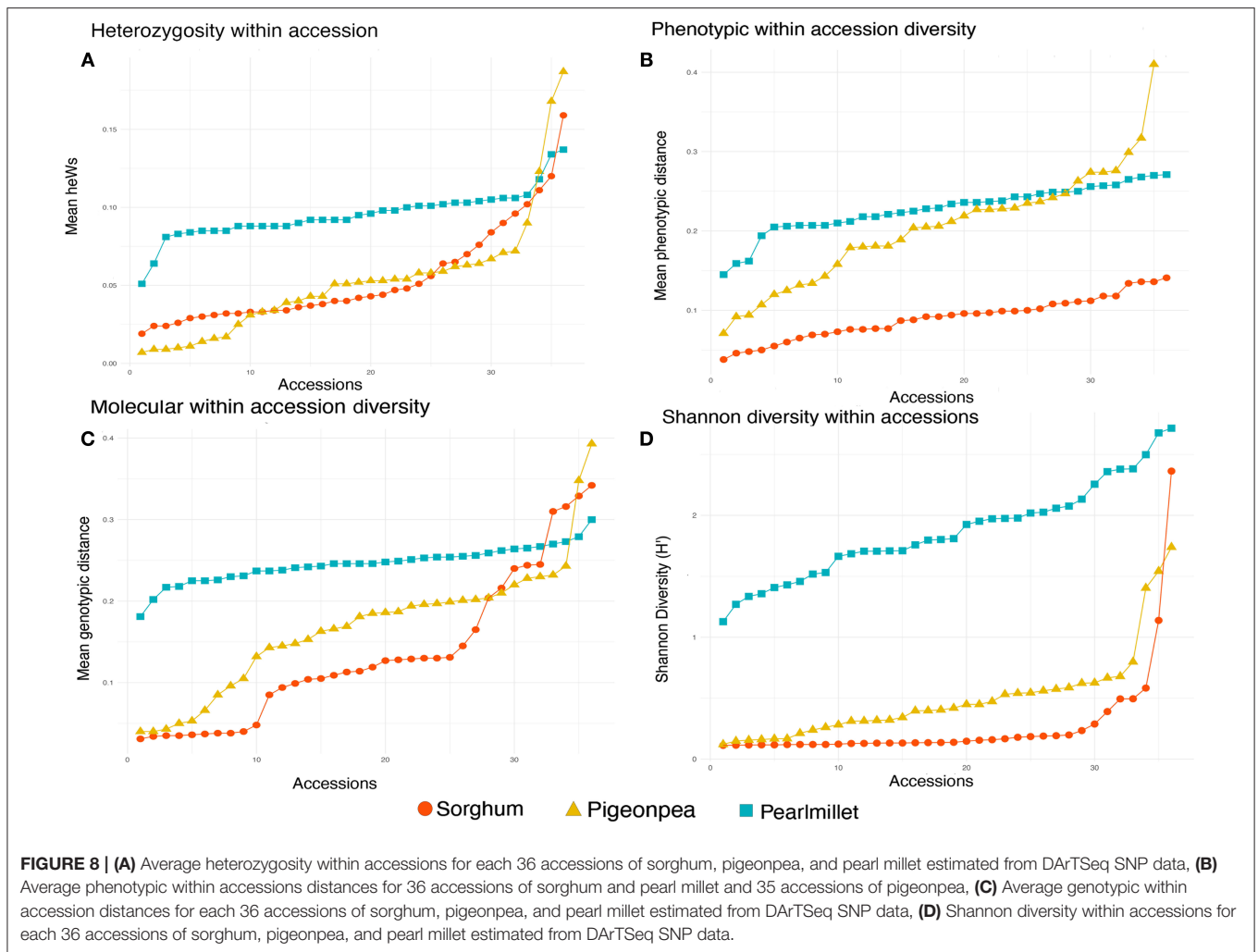
Population structure analyses indicated that most of the accessions in sorghum were uniform enough to cluster individuals of single accession together as a singleton clusters. In sorghum, accessions IS 29508 and IS 29605 were clustered into a single population in both DAPC and ward.D2 clustering with genotypic distances, indicating the presence of high similarity and common alleles in these two accessions. The accession IS 33844 showed high uniformity with a low within accession diversity (0.031), and a selection from this landrace has been released as a variety in India as ‘Parbhani moti’ (Upadhyaya and Vetriventhan, 2018). Pearl millet and pigeonpea showed a higher heterogeneity within accessions, while most of the accessions



showed mixed populations. In pigeonpea and pearl millet only some accessions showed singleton clustering. High population mixtures in these crops correspond to their pollination behavior, and sharing of alleles between populations. Landraces generally differ between populations, based on the intensity of selection imposed by farmers, their pollination mechanisms, the level of gene-flow within and between population, and level of exchange of seed materials between farmers. Previously several authors reported pollen flow between populations and the mixing of landrace populations in sorghum and pigeonpea (Songok et al., 2010; Kassa et al., 2012; Adugna, 2014; Westengen et al., 2014). Harlan (1965) reported the gene-flow from weeds to landraces and several other authors (Ellstrand et al., 1999; Jarvis and Hodgkin, 1999; Messeguer, 2003; Gompert and Buerkle, 2016) reported the transfer of genes into landraces from various sources in both self and outcrossing species. Also, some studies reported the mixing of the population by a considerable exchange of seeds within cultures or regions (Louette, 1997). The level of heterogeneity and diversity in landraces are crop-specific and associated with their mode of fertilization (Villa et al., 2005) and also several authors (Hammer et al., 1996; Zeven, 1998) stated the influence of mode of pollination in various population

genetics factors over the course of evolution of landraces. Hence, complying to the effect of mode of reproduction on diversity, a higher degree of outcrossing (about 85%) in pearl millet (Burton, 1983) could impose a higher diversity in pearl millet, in comparison to lower outcrossing crops such as sorghum (about 18%) (Barnaud et al., 2008) and an intermediate outcrossing crops (about 30%) like pigeonpea (Saxena et al., 1990), and this varies with species.

Most of the accessions that showed relatively higher within genetic distances in sorghum and pearl millet were wild accessions. Thus, using wild accessions in this study helped us in the comparative assessment with landraces and also aided in the better understanding of the effect of domestication and different evolutionary forces that shaped the landraces. Historically farmers conserving landraces on-farm and multiplied desirable phenotypes, which survived both natural and artificial selection. The effect of this farmers' selection led to local adaptations and variations within the landrace populations (Zeven, 1998). Teshome et al. (2016) studied the maintenance of landrace diversity in sorghum by farmers belonging to different regions in Ethiopia and reported a narrow preference to specific economic traits and selection by farmers. Thus, the wild accessions in this



case lack of farmers' selections and its obligatory to be highly diverse as these are evolving under natural selection.

Comparing diversity of the three crops in our study, heterozygosity (**Figure 8A**), phenotypic (**Figure 8B**), and genotypic (**Figure 8C**) within accession diversity of sorghum were notably low for most of its accessions, intermediate for most of the accessions of pigeonpea and followed a more stable trend around the maximal values for pearl millet. Similar to the molecular within accession distances, Shannon diversity revealed diversity estimates, scaled over the higher values for pearl millet, followed by an intermediate in pigeonpea and lower estimates for sorghum (**Figure 8D**). However, in sorghum and pigeonpea both highly diverse and highly uniform accessions with maximal and minimal estimates of genotypic distances and Shannon diversity were observed. The higher diversity estimates indicate the presence of higher variability within accessions. In case of pigeonpea most of the accessions were found to have molecular within accession distances <0.250 except two accessions *viz.*, ICP 10880 (0.348) and ICP 10889 (0.393). On further investigation into the individual plant within accession distances of these accessions, it appeared that,

some individuals within these accessions were diverse from all the other individuals of the respective accession. Such that, the accession ICP 10880 had two individuals that were highly divergent from all other individuals by a mean distance of 0.410 and 0.434. Also these individuals were found to cluster separately in hierarchal clustering. Same for the accession ICP 10889, where some individuals were highly divergent from the other. In case of sorghum, most of the accessions had a molecular within accession distances <0.250 except three wild accessions *viz.*, IS 10897 (0.316), IS 14485 (0.329), IS 18833 (0.342) and one landrace IS 27325 (0.310). In the landrace IS 27325, it can be seen that the individuals are divided into three subgroups in hierarchal clustering. Thus, higher diversity in some landraces of sorghum and pigeonpea can be due their pollination behavior, which ultimately influences the population substructure. The lower outcrossing in these crops offers the higher probability of fixation of various alleles within a fewer members or individuals, restricting the frequency/occurrence of some allele within a small group of a landrace population, thus gradually over generations, forming distinct subpopulations within groups. These varied groups of individuals are however

not phenotypically variable enough to consider it as separate population, but however assimilated a genetically distinct fingerprint from various elements throughout the course of evolution. Similar cases of extreme values of low and high diversity were previously encountered by researchers. Zeven (1998) emphasized the low diversity and increased homozygosity in inbreeding accessions and also explained the influence of farmers' selection and sampling strategies for reduction of diversity in landrace populations. Adugna (2014) and Westengen et al. (2014) found both high and low within-population diversity in sorghum landraces cultivated in Ethiopia and reasoned the low within landrace diversity could be due to farmers' sampling during migration, as farmers tend to carry few heads during migration and settlements.

Based on the level of diversity within each accession of different crops, appropriate conservation and regeneration strategy should be followed to conserve the genetic integrity and diversity of landraces. ICRISAT genebank follows various pollination control and sampling strategies to maintain the genetic integrity and diversity within accessions, while regenerating different crops. Theoretically, selfing will be a good strategy to maintain the genetic integrity and diversity in self-pollinated crops and often-cross pollinated crops (out-crossing >5%), because of the low effect of inbreeding depression, and to preserve alleles within the population. In cross-pollinated species like pearl millet, sib mating is the best strategy to mimic the random mating, and for that ICRISAT genebank performs cluster bagging (bagging few panicles of different individuals of the same accession) that reduces the effect of inbreeding depression. However, in both cases, the appropriate population size needs to be ensured while regeneration for capturing the rare alleles. Small sample sizes while regenerating landraces may lead to genetic drift which results in the loss of some rare alleles. Crossa (1989) based on his results on stimulated populations, reported a practical system for maize regeneration, wherein the author discussed that the ideal system of regeneration involves equalizing the genetic contribution of parents and avoiding small population sizes and, also Crossa (1995) suggested a practical seed sample size of 130–200 in monoecious crops for retaining the rare alleles in most of the loci. FAO standards specify a sample size of 30 individuals in a completely random mating population and 60 individuals for completely selfing species to capture 95% of the alleles which have a frequency >0.05 (FAO, 2014). However, in sorghum, pigeonpea, and pearl millet, no detailed molecular studies were done previously utilizing NGS tools to determine optimum population size requirements for regeneration. Therefore, we estimated the minimum sample size to capture 95% of the SNP alleles spread throughout the whole genome with an expected probability of 95% based on the least frequent allele or the frequency of the rarest allele for each accession following Crossa (1989). From our study, seed sample sizes were found to be minimal for sorghum (47–101), and pigeonpea (77–89), and high for pearl millet (155–203). The sample size required to conserve the genetic integrity of germplasm depends largely on the frequency of the least common alleles or genotypes.

In conclusion, sorghum, pigeonpea, and pearl millet accessions showed higher within and among accession diversity, indicating that the regeneration strategies at ICRISAT genebank are appropriate to ensure the genetic integrity of each accession. Information from this study will support genebank curators in understanding within accession variability and assists in devising scientific sampling strategies (sample size) for regeneration to maintain the genetic integrity and variability. This could also help breeders in the utilization end to understand the population dynamics and subpopulation structure, to forward the material with appropriate breeding techniques.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: <http://dataverse.icrisat.org/dataset.xhtml?persistentId=doi:10.21421/D2/CCSOZ8> for pigeonpea <http://dataverse.icrisat.org/dataset.xhtml?persistentId=doi:10.21421/D2/WU4JFA> for pearl <http://dataverse.icrisat.org/dataset.xhtml?persistentId=doi:10.21421/D2/DSYLHB> for sorghum.

AUTHOR CONTRIBUTIONS

VCRA, MV, SD, and AR contributed to conception and design of the study. VCRA, MV, VA, RS, VK, PS, and SR conducted field experiments and data collection. This work is part of VA's thesis research. SG supported student research as chairman. VA and MV performed the statistical analysis and wrote the first draft of the manuscript. VCRA reviewed and approved the first draft. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

We acknowledge funds by global diversity crop trust (GCDT) and genebank platform for funding support for this study.

ACKNOWLEDGMENTS

We thank Dr. Juan Andres (Biometrics and Statistics Unit Head, CIMMYT, Mexico) for his kind help in AMOVA for the DArTSeq data and Shannon diversity. Also, we extend our thanks to Dr. Jorge Franco (Professor of Biometrics and Statistics, Universidad de la República, Uruguay), Dr. Carolina Paola Sansaloni (High-throughput genotyping and sequencing specialist, CIMMYT, Mexico) and Dr. Fernando Henrique Ribeiro (Associate Scientist-Agricultural Statistics, CIMMYT, Mexico) for their kind support in calculation of genotypic diversity.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.587426/full#supplementary-material>

Supplementary Figure 1 | Cluster dendrogram with unbiased bootstrap probability values for edges, with ward.D2 clustering for Gower's distances, for single plant phenotypic data **(A)** The cluster dendrogram of sorghum, **(B)** the cluster dendrogram of pigeonpea, and **(C)** Cluster dendrogram of pearl millet.

Supplementary Figure 2 | Cluster dendrogram with bootstrap probability values for edges, with ward.D2 clustering for Modified Roger's distances, for single plant genotypic data **(A)** The cluster dendrogram of sorghum, **(B)** the cluster dendrogram of pigeonpea, and **(C)** Cluster dendrogram of pearl millet.

Supplementary Figure 3 | Values of BIC vs. number of clusters with maximum of 70 clusters in DAPC analysis for **(A)** sorghum, **(B)** pigeonpea, and **(C)** pearl millet.

Supplementary Figure 4 | Population structure using the posterior membership probabilities, using $K = 36$ in DAPC analysis: **(A)** The population structure of sorghum. **(B)** Population structure of pigeonpea and **(C)** Population structure of pearl millet.

Supplementary Table 1 | List of accessions of sorghum used for phenotypic and genotypic within accession diversity evaluation.

Supplementary Table 2 | List of accessions of pigeonpea used for phenotypic and genotypic within accession diversity evaluation.

Supplementary Table 3 | List of accessions of pearl millet used for phenotypic and genotypic within accession diversity evaluation.

Supplementary Table 4 | List of the quantitative and qualitative traits recorded in sorghum, pearl millet and pigeonpea.

Supplementary Table 5 | Mean grouping by Student-Newman-Keuls Test of sorghum, pearl millet and pigeonpea for all the quantitative traits recorded.

Supplementary Table 6 | Levene's test for significant differences in population variances.

Supplementary Table 7 | Bootstrapping values (Jaccard coefficient) of clusters for phenotypic data.

Supplementary Table 8 | Number of DArTSeq-SNPs in each chromosome of sorghum, pearl millet and pigeonpea after filtering.

Supplementary Table 9 | Bootstrapping (Jaccard coefficients) values of clusters for DArTSeq SNP data.

Supplementary Table 10 | Probability of attribution of all the single plants in each accession into different groups based on the discriminant analysis of principle components on DArTSeq-SNP data ($K = 36$).

Supplementary Table 11 | Number of rare alleles preserved in the recommended sample size estimated from DArTSeq-SNP data of sorghum, pigeonpea, and pearl millet.

REFERENCES

- Adugna, A. (2014). Analysis of *in situ* diversity and population structure in Ethiopian cultivated *Sorghum bicolor* (L.) landraces using phenotypic traits and SSR markers. *Springerplus* 3:212. doi: 10.1186/2193-1801-3-212
- Al Khanjari, S., Hammer, K., Buerkert, A., and Röder, M. S. (2007). Molecular diversity of Omani wheat revealed by microsatellites: II. hexaploid landraces. *Genet. Resour. Crop Evol.* 54, 1407–1417. doi: 10.1007/s10722-006-9125-1
- Animasaun, D., Morakinyo, J., Mustapha, O., and Krishnamurthy, R. (2015). Assessment of genetic diversity in accessions of pearl millet (*Pennisetum glaucum*) and napier grass (*Pennisetum purpureum*) using microsatellite (ISSR) markers. *Iran. J. Genet. Plant Breed.* 4, 25–35.
- Astley, D. (1991). 2. Exploration: methods and problems of exploration and field collecting. *Biol. J. Linn. Soc.* 43, 11–22. doi: 10.1111/j.1095-8312.1991.tb00579.x
- Barilli, E., Cobos, M. J., Carrillo, E., Kilian, A., Carling, J., and Rubiales, D. (2018). A high-density integrated DArTseq SNP-based genetic map of *pisum fulvum* and identification of QTLs controlling rust resistance. *Front. Plant Sci.* 9:167. doi: 10.3389/fpls.2018.00167
- Barnaud, A., Trigueros, G., McKey, D., and Joly, H. I. (2008). High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? *Heredity* 101, 445–452. doi: 10.1038/hdy.2008.77
- Bashir, E. M. A., Ali, A. M., Ali, A. M., Mohamed, E. T. I., Melchinger, A. E., Parzies, H. K., et al. (2015). Genetic diversity of sudanese pearl millet (*Pennisetum glaucum* (L.) R. Br.) landraces as revealed by SSR markers, and relationship between genetic and agro-morphological diversity. *Genet. Resour. Crop Evol.* 62, 579–591. doi: 10.1007/s10722-014-0183-5
- Bellon, M. R. (2009). “Do we need crop landraces for the future? Realizing the global option value of *in situ* conservation,” in *Agrobiodiversity and Economic Development*, 1st Edn., eds. A. Kontoleon, U. Pasqual, and M. Smale (Abington: Routledge).
- Bhattacharjee, R., Bramel, P., Hash, C., Kolesnikova-Allen, M., and Khairwal, I. (2002). Assessment of genetic diversity within and between pearl millet landraces. *Theor. Appl. Genet.* 105, 666–673. doi: 10.1007/s00122-002-0917-1
- Brown, A. H. D. (1978). Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* 52, 145–157. doi: 10.1007/BF00282571
- Budak, H., Pedraza, F., Cregan, P. B., Baenziger, P. S., and Dweikat, I. (2003). Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of pearl millet germplasm. *Crop Sci.* 43, 2284–2290. doi: 10.2135/cropsci2003.2284
- Burton, G. W. (1983). “Breeding pearl millet,” in *Plant Breeding Reviews-Volume 1*, ed J. Janick (Boston, MA: Springer US), 162–182. doi: 10.1007/978-1-4684-8896-8_6
- Busso, C. S., Devos, K. M., Ross, G., Mortimore, M., Adams, W. M., Ambrose, M. J., et al. (2000). Genetic diversity within and among landraces of pearl millet (*Pennisetum glaucum*) under farmer management in West Africa. *Genet. Resour. Crop Evol.* 47, 561–568. doi: 10.1023/A:1008767220320
- Chandra-Shekara, A. C., Prasanna, B. M., Bhat, S. R., and Singh, B. B. (2007). Genetic diversity analysis of elite pearl millet inbred lines using RAPD and SSR markers. *J. Plant Biochem. Biotechnol.* 16, 23–28. doi: 10.1007/BF03321924
- Chowdari, K. V., Davierwala, A. P., Gupta, V. S., Ranjekar, P. K., and Govila, O. P. (1998). Genotype identification and assessment of genetic relationships in pearl millet [*Pennisetum glaucum* (L.) R. Br.] using microsatellites and RAPDs. *Theor. Appl. Genet.* 97, 154–162. doi: 10.1007/s001220050880
- Crossa, J. (1989). Methodologies for estimating the sample size required for genetic conservation of outbreeding crops. *Theor. Appl. Genet.* 77, 153–161. doi: 10.1007/BF00266180
- Crossa, J. (1995). “Sample size and effective population size in seed regeneration of monococious species,” in *Regeneration of Seed Crops and Their Wild Relatives: Proceedings of a Consultation Meeting* (ICRISAT) (Hyderabad: Bioversity International), 140.
- D'Andrea, A. C., and Casey, J. (2002). Pearl millet and Kintampo subsistence. *Afr. Archaeol. Rev.* 19, 147–173. doi: 10.1023/A:1016518919072
- de Mendiburu F. (2013). *Statistical Procedures for Agricultural Research. Package 'Agricolae,' Version 1.4-4*. Vienna: Comprehensive R Archive Network, Institute for Statistics and Mathematics. Available online at: <http://cran.r-project.org/web/packages/agricolae/agricolae.pdf>
- Djè, Y., Forcioli, D., Ater, M., Lefebvre, C., and Vekemans, X. (1999). Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. *Theor. Appl. Genet.* 99, 157–163. doi: 10.1007/s001220051220
- Dray, S., Dufour, A.-B., and Chessel, D. (2007). The {ade4} package – [II]: two-table and {K}-table methods. *R News* 7, 47–52.
- Dreisigacker, S., Zhang, P., Warburton, M. L., Skovmand, B., Hoisington, D., and Melchinger, A. E. (2005). Genetic diversity among and within CIMMYT wheat landrace accessions investigated with SSRs and implications for plant genetic resources management. *Crop Sci.* 45, 653–661. doi: 10.2135/cropsci2005.0653
- Dwivedi, S. L., Ceccarelli, S., Blair, M. W., Upadhyaya, H. D., Are, A. K., and Ortiz, R. (2016). Landrace germplasm for improving yield and abiotic stress adaptation. *Trends Plant Sci.* 21, 31–42. doi: 10.1016/j.tplants.2015.10.012

- Edet, O. U., Gorafi, Y. S. A., Nasuda, S., and Tsujimoto, H. (2018). DArTseq-based analysis of genomic relationships among species of tribe Triticeae. *Sci. Rep.* 8:16397. doi: 10.1038/s41598-018-34811-y
- Ellstrand, N. C., Prentice, H. C., and Hancock, J. F. (1999). Domesticated plants into their wild relatives. *Annu. Rev. Ecol. Syst.* 30, 539–563. doi: 10.1146/annurev.ecolsys.30.1.539
- Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131, 479–491.
- FAO (2014). *Genebank Standards for Plant Genetic Resources for Food and Agriculture*. Rome: Food and Agriculture Organization.
- Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression*. 3rd ed. Thousand Oaks, CA: Sage Publications.
- Frankel, O., and Soule, M. E. (1981). *Conservation and Evolution*. 1st Edn. Cambridge: Cambridge University Press.
- Gompert, Z., and Buerkle, C. A. (2016). What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evol. Appl.* 9, 909–923. doi: 10.1111/eva.12380
- Goodman, M. M., and Stuber, C. W. (1983). Races of maize. 6: Isozyme variation among races of maize in Bolivia. *Maydica* 28, 169–188.
- Govindaraj, M., Selvi, B., Prabhu, D. A., and Rajarathinam, S. (2009). Genetic diversity analysis of pearl millet (*Pennisetum glaucum* [L.] R. Br.) accessions using molecular markers. *Afr. J. Biotechnol.* 8, 6046–6052. doi: 10.5897/AJB2009.000-9493
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27, 857–871. doi: 10.2307/2528823
- Hagenblad, J., Zie, J., and Leino, M. W. (2012). Exploring the population genetics of genebank and historical landrace varieties. *Genet. Resour. Crop Evol.* 59, 1185–1199. doi: 10.1007/s10722-011-9754-x
- Hammer, K., Knüpfer, H., Xhuvli, L., and Perrino, P. (1996). Estimating genetic erosion in landraces - two case studies. *Genet. Resour. Crop Evol.* 43, 329–336. doi: 10.1007/BF00132952
- Harlan, J. R. (1965). The possible role of weed races in the evolution of cultivated plants. *Euphytica* 14, 173–176. doi: 10.1007/BF00038984
- Harlan, J. R. (1971). Agricultural origins: centers and noncenters. *Science* 174, 468–474. doi: 10.1126/science.174.4008.468
- Harlan, J. R., and de Wet, J. M. J. (1972). A simplified classification of cultivated sorghum. *Crop Sci.* 12, 172–176. doi: 10.2135/cropsci1972.0011183X001200020005x
- Hawkes, J. G. (1983). *The Diversity of Crop Plants*. Cambridge: Harvard University Press.
- Hennig, C. (2020). *fpc: Flexible Procedures for Clustering*. Available online at: <https://cran.r-project.org/package=fpc>
- IBPGR and ICRISAT (1993a). *Descriptors for Pearl Millet [Pennisetum glaucum (L.) R. Br.]*. Rome; Patancheru: Plant Genetic Resources; International Crops Research Institute for the Semi-Arid Tropics.
- IBPGR and ICRISAT (1993b). *Descriptors for Pigeonpea [Cajanus cajan (L.) Millsp.]*. Rome; Patancheru: Plant Genetic Resources; International Crops Research Institute for the Semi-Arid Tropics.
- IBPGR and ICRISAT (1993c). *Descriptors for Sorghum [Sorghum bicolor (L.) Moench]*. Rome; Patancheru: Plant Genetic Resources; International Crops Research Institute for the Semi-Arid Tropics.
- Jarvis, D. I., and Hodgkin, T. (1999). Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. *Mol. Ecol.* 8, S159–S173. doi: 10.1046/j.1365-294X.1999.00799.x
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Jones, H., Leigh, F. J., Mackay, I., Bower, M. A., Smith, L. M. J., Charles, M. P., et al. (2008). Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the fertile crescent. *Mol. Biol. Evol.* 25, 2211–2219. doi: 10.1093/molbev/msn167
- Kamvar, Z. N., Brooks, J. C., and Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front. Genet.* 6:208. doi: 10.3389/fgene.2015.00208
- Kassa, M. T., Penmetsa, R. V., Carrasquilla-Garcia, N., Sarma, B. K., Datta, S., Upadhyaya, H. D., et al. (2012). Genetic patterns of domestication in pigeonpea (*Cajanus cajan* (L.) Millsp.) and wild cajan relatives. *PLoS ONE* 7:e39563. doi: 10.1371/journal.pone.0039563
- Keuls, M. (1952). The use of the studentized range in connection with an analysis of variance. *Euphytica* 1, 112–122. doi: 10.1007/BF01908269
- Kumar, A., Arya, L., Kumar, V., and Sharma, S. (2006). Inter simple sequence repeat (ISSR) analysis of cytoplasmic male sterile, male fertile lines and hybrids of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Crop Sci.* 1, 117–119.
- Kyratzis, A. C., Nikoloudakis, N., and Katsiotis, A. (2019). Genetic variability in landraces populations and the risk to lose genetic variation. the example of landrace “Kyperounda” and its implications for ex situ conservation. *PLoS ONE* 14:e0224255. doi: 10.1371/journal.pone.0224255
- Levene, H. (1960). “Robust tests for the equality of variance,” in *Contributions to Probability and Statistics*, ed I. Olkin (Palo Alto, CA: Stanford University Press), 278–292.
- Louette, D. (1997). “Seed exchange among farmers and gene flow among maize varieties in traditional agricultural systems,” in *Gene Flow Among Maize Landraces, Improved Maize Varieties, and Teosinte: Implications for Transgenic Maize* (Mexico: CIMMYT), 56–66.
- Mace, E. S., Buhariwalla, K. K., Buhariwalla, H. K., and Crouch, J. H. (2003). A high-throughput DNA extraction protocol for tropical molecular breeding programs. *Plant Mol. Biol. Rep.* 21, 459–460. doi: 10.1007/BF02727596
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2019). *Package “Cluster: Cluster Analysis Basics and Extensions”, Version 2.0.9, 1–82*. Available online at: <https://svn.r-project.org/R-packages/trunk/cluster>
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.
- Martin, G. B., and Adams, M. W. (1987). Landraces of *Phaseolus vulgaris* (Fabaceae) in Northern Malawi. I. regional variation. *Econ. Bot.* 41, 190–203. doi: 10.1007/BF02858965
- Messeguer, J. (2003). Gene flow assessment in transgenic plants. *Plant Cell. Tissue Organ Cult.* 73, 201–212. doi: 10.1023/A:1023007606621
- Michaelis, A., Green, M., and Rieger, R. (1991). *Glossary of Genetics: Classical and Molecular*. Berlin: Springer Science & Business Media.
- Murtagh, F., and Legendre, P. (2014). Ward’s hierarchical agglomerative clustering method: which algorithms implement Ward’s criterion? *J. Classif.* 31, 274–295. doi: 10.1007/s00357-014-9161-z
- Ndjondjop, M. N., Semagn, K., Sow, M., Manneh, B., Gouda, A. C., Kpeki, S. B., et al. (2018). Assessment of genetic variation and population structure of diverse rice genotypes adapted to lowland and upland ecologies in Africa using SNPs. *Front. Plant Sci.* 9:446. doi: 10.3389/fpls.2018.00446
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.* 70, 3321–3323. doi: 10.1073/pnas.70.12.3321
- Newman, D. (1939). The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika* 31, 20–30. doi: 10.1093/biomet/31.1-2.20
- Pailles, Y., Ho, S., Pires, I. S., Tester, M., Negrão, S., and Schmöcke, S. M. (2017). Genetic diversity and population structure of two tomato species from the galapagos islands. *Front. Plant Sci.* 8:138. doi: 10.3389/fpls.2017.00138
- Pattanashetti, S. K., Upadhyaya, H. D., Dwivedi, S. L., Vetriventhan, M., and Reddy, K. N. (2016). “Pearl millet,” in *Genetic and Genomic Resources for Grain Cereals Improvement*, eds M. Singh and H. D. Upadhyaya (San Diego, CA: Academic Press), 253–289. doi: 10.1016/B978-0-12-802000-5.00006-X
- Pressoir, G., and Berthaud, J. (2004). Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity* 92, 88–94. doi: 10.1038/sj.hdy.6800387
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Available online at: <https://www.r-project.org/>
- Raman, R., Qiu, Y., Coombes, N., Song, J., Kilian, A., and Raman, H. (2017). Molecular diversity analysis and genetic mapping of pod shatter resistance loci in *Brassica carinata* L. *Front. Plant Sci.* 8:1765. doi: 10.3389/fpls.2017.01765
- Ramanatha Rao, V., and Hodgkin, T. (2002). Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell. Tissue Organ Cult.* 68, 1–19. doi: 10.1023/A:1013359015812
- Saxena, K. B., Singh, L., and Gupta, M. D. (1990). Variation for natural out-crossing in pigeonpea. *Euphytica* 46, 143–148. doi: 10.1007/BF00022307
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

- Shewayrga, H., Jordan, D. R., and Godwin, I. D. (2008). Genetic erosion and changes in distribution of sorghum (*Sorghum bicolor* L. (Moench)) landraces in north-eastern Ethiopia. *Plant Genet. Resour.* 6, 1–10. doi: 10.1017/S1479262108923789
- Songok, S., Ferguson, M., Muigai, A. W., and Silim, S. (2010). Genetic diversity in pigeonpea [*Cajanus cajan* (L.) Millsp.] landraces as revealed by simple sequence repeat markers. *Afr. J. Biotechnol.* 9, 3231–3241.
- Suzuki, R., Terada, Y., and Shimodaira, H. (2019). *pvclust: Hierarchical Clustering With P-Values via Multiscale Bootstrap Resampling*. Available online at: <https://cran.r-project.org/package=pvclust>
- Teshome, A. A., Fahrig, L., Torrance, J. K., Lambert, J. D., and Arnason, T. J. (2016). Maintenance of sorghum (*Sorghum bicolor*, *Poaceae*) landrace diversity by farmers' selection in Ethiopia. *Econ. Botany* 53, 79–88. doi: 10.1007/BF02860796
- Tostain, S., and Marchais, L. (1989). Enzyme diversity in pearl millet (*Pennisetum glaucum*). *Theor. Appl. Genet.* 77, 634–640. doi: 10.1007/BF00261235
- Tostain, S., Riandey, M.-F., and Marchais, L. (1987). Enzyme diversity in pearl millet (*Pennisetum glaucum*). *Theor. Appl. Genet.* 74, 188–193. doi: 10.1007/BF00289967
- Upadhyaya, H. D., and Vetriventhan, M. (2018). “Ensuring the genetic diversity of sorghum,” in *Achieving Sustainable Cultivation of Sorghum*, ed W. Rooney (Cambridge: Burleigh Dodds Science Publishing Limited), 546. doi: 10.19103/AS.2017.0015.06
- Villa, T. C. C., Maxted, N., Scholten, M., and Ford-Lloyd, B. (2005). Defining and identifying crop landraces. *Plant Genet. Resour.* 3, 373–384. doi: 10.1079/PGR200591
- Westengen, O. T., Okongo, M. A., Onok, L., Berg, T., Upadhyaya, H., Birkeland, S., et al. (2014). Ethnolinguistic structuring of sorghum genetic diversity in Africa and the role of local seed systems. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14100–14105. doi: 10.1073/pnas.1401646111
- Wright, S. (1978). “Variability within and among natural populations,” in *Evolution and the Genetics of Populations: A Treatise in Four Volumes*, Vol. 1 (Chicago, IL: University of Chicago Press), 91–96.
- Xie, X. M., Zhou, F., Zhang, X. Q., and Zhang, J. M. (2010). Genetic variability and relationship between MT-1 elephant grass and closely related cultivars assessed by SRAP markers. *J. Genet.* 88, 281–290. doi: 10.1007/s12041-009-0041-y
- Zeven, A. C. (1998). Landraces: a review of definitions and classifications. *Euphytica* 104, 127–139. doi: 10.1023/A:1018683119237

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Citation: Allan V, Vetriventhan M, Senthil R, Geetha S, Deshpande S, Rathore A, Kumar V, Singh P, Reddymalla S and Azevedo VCR (2020) Genome-Wide DArTSeq Genotyping and Phenotypic Based Assessment of Within and Among Accessions Diversity and Effective Sample Size in the Diverse Sorghum, Pearl Millet, and Pigeonpea Landraces. *Front. Plant Sci.* 11:587426. doi: 10.3389/fpls.2020.587426

Copyright © 2020 Allan, Vetriventhan, Senthil, Geetha, Deshpande, Rathore, Kumar, Singh, Reddymalla and Azevedo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Breeding Strategies and Challenges in the Improvement of Blast Disease Resistance in Finger Millet. A Current Review

Wilton Mbinda^{1,2*} and Hosea Masaki¹

¹ Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya, ² Pwani University Biosciences Research Centre (PUBReC), Pwani University, Kilifi, Kenya

OPEN ACCESS

Edited by:

Sukhjiwan Kaur,
Agriculture Victoria, Australia

Reviewed by:

Mohammad Ismail,
Indian Agricultural Research Institute
(ICAR), India

Benoit Bertrand,
Institut National de la Recherche
Agronomique (INRA), Centre
de Coopération Internationale en
Recherche Agronomique pour le
Développement, France
Prabhakaran Sambasivam,
Griffith University, Australia

*Correspondence:

Wilton Mbinda
w.mbinda@pu.ac.ke;
wilton.mbinda@gmail.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 04 September 2020

Accepted: 08 December 2020

Published: 08 January 2021

Citation:

Mbinda W and Masaki H (2021)
Breeding Strategies and Challenges
in the Improvement of Blast Disease
Resistance in Finger Millet. A Current
Review. *Front. Plant Sci.* 11:602882.
doi: 10.3389/fpls.2020.602882

Climate change has significantly altered the biodiversity of crop pests and pathogens, posing a major challenge to sustainable crop production. At the same time, with the increasing global population, there is growing pressure on plant breeders to secure the projected food demand by improving the prevailing yield of major food crops. Finger millet is an important cereal crop in southern Asia and eastern Africa, with excellent nutraceutical properties, long storage period, and a unique ability to grow under arid and semi-arid environmental conditions. Finger millet blast disease caused by the filamentous ascomycetous fungus *Magnaporthe oryzae* is the most devastating disease affecting the growth and yield of this crop in all its growing regions. The frequent breakdown of blast resistance because of the susceptibility to rapidly evolving virulent genes of the pathogen causes yield instability in all finger millet-growing areas. The deployment of novel and efficient strategies that provide dynamic and durable resistance against many biotypes of the pathogen and across a wide range of agro-ecological zones guarantees future sustainable production of finger millet. Here, we analyze the breeding strategies currently being used for improving resistance to disease and discuss potential future directions toward the development of new blast-resistant finger millet varieties, providing a comprehensive understanding of promising concepts for finger millet breeding. The review also includes empirical examples of how advanced molecular tools have been used in breeding durably blast-resistant cultivars. The techniques highlighted are cost-effective high-throughput methods that strongly reduce the generation cycle and accelerate both breeding and research programs, providing an alternative to conventional breeding methods for rapid introgression of disease resistance genes into favorable, susceptible cultivars. New information and knowledge gathered here will undoubtedly offer new insights into sustainable finger millet disease control and efficient optimization of the crop's productivity.

Keywords: allele mining, finger millet, gene pyramiding, *M. oryzae*, marker assisted selection, molecular breeding, QTL mapping, transgenesis

INTRODUCTION

Environmental stresses cause reduced growth and significant yield losses in food crops. Avenues for improvement of crops to combat individual and multiple stresses are major breeding goals, and different approaches have been used to improve stress tolerance (Calanca, 2017). Among the biotic stresses, pests and diseases are the most important limiting factors that affect finger millet production worldwide. Finger millet blast caused by the filamentous fungus *Magnaporthe oryzae* (sexual amorph *Pyricularia oryzae*) is the most devastating disease affecting the production and productivity of finger millet because of its destructive nature under favorable conditions and its wide distribution in all finger millet-growing areas (Yoshida et al., 2016). *M. oryzae* infects a finger millet plant during nearly all growth stages and reduces the crop grain yield by up to 100% (Senthil et al., 2012; Takan et al., 2012). Due to the importance of finger millet blast disease, effective disease control measures are needed to ensure global food security, especially in arid and semi-arid regions of African and Asia where the crop is majorly cultivated. Over the years, various pursuits have been made to develop new cultivars that are resistant to the disease. The unavailability of the whole-genome sequence and the limited genomic resources of finger millet has greatly hampered studies of the genetics of resistance to the blast disease compared with other major cereals. As a result of this shortcoming, the understanding of broad-spectrum resistance to finger millet blast disease remains a knowledge gap. Previous work on the pathosystem of the blast fungus on finger millet relied primarily on the phenotypic features and virulence tests using various hosts (Wu et al., 2014; Que et al., 2019; Rasool et al., 2020). These studies concentrated on screening and selection of finger millet cultivars or new advanced lines toward selected strains of blast fungi but with limited success, because the phenotypic traits obtained are extremely variable due to the genetic instability of the blast pathogen. In addition, it often takes a long time and high cost for breeders to develop new varieties with broad-spectrum resistance to blast disease when these strategies are adopted. Additionally, the studies are influenced by environmental pressures and are prone to human errors, leading to ambiguous results.

With the advent of new biotechnological tools, current research strategically focuses on understanding the biotic stresses on finger millet, particularly in advancing the molecular genetics of blast disease in order to develop an integrated management system for blast disease resistance in finger millet. With the advancement of high-throughput sequencing platforms, there has been a tremendous increase in the modern genomic tools available, such as molecular markers, expressed sequence tags (ESTs), gene expression profiling, genome-wide association studies, genetic transformations, and, more recently, genome editing, have been used successfully in various crops to explore the genetic basis of stress tolerance for guidance in the development of plants of superior quality. DNA molecular markers have also been used for population genetics and evaluation of genetic variations that occur between and within plant populations, and their polymorphic structure and level can

be invaluable in crop breeding. This review outlines a set of recent molecular and genomic tools that are used to study the finger millet blast fungus.

Finger Millet and Finger Millet Blast Importance

Finger millet [*Eleusine coracana* (L.) Gaertn.] is an allotetraploid ($2n = 4X = 36$) member of the Poaceae family. The crop is mostly grown and consumed by people in the poverty-stricken arid and semi-arid tropics of Asia and sub-Saharan Africa (Takan et al., 2012). Its grains have excellent nutraceutical properties, such as high dietary fiber content, amino acids (methionine, phenylalanine, tryptophan, cysteine, isoleucine, and leucine), vitamin B complex, calcium, and iron compared with maize, rice, wheat, and sorghum (Kumar et al., 2016b; Gupta et al., 2017). It is also gluten-free and can be stored for a long period (Chandrasekara and Shahidi, 2010). As a member of the small millets, finger millet is the most climate-resilient crop which can be cultivated under a diverse range of climatic conditions. It ranks fourth on a global scale of production, followed by sorghum, pearl millet, and foxtail millet (Upadhyaya et al., 2007). These qualities make the finger millet an important food and nutritional security crop and a valuable genomic resource. Although the global finger millet production has been increasing, reaching 4.5 million tons in 2018 (FAOSTAT, 2019), the increase is not concomitant with the demand for finger millet because of the rapidly increasing human population and industrialization. To date, no published data exists on associated economic loss due blast disease in finger millet. To overcome this challenge, there is a need to increase finger millet production by at least 40%, like other major cereals (Kumar et al., 2018).

The availability and access to diverse genetic resources is central to genetic improvement of any crop. These genetic resources have to be characterized for their effective utilization in crop improvement programs. Field and *in vitro* genebanks constitute a huge pool of finger millet germplasm collections. As of 2010, 35382 finger millet accessions were conserved in gene-banks across the world. The National Bureau of Plant Genetic Resources (India) and the International Crops Research Institute for the Semi-Arid (ICRISAT) genebanks across the world accounted for 26.9% and 16.8% of the global collections. Other institutions including Kenya Agricultural & Livestock Research Organization (Kenya), National Agricultural Research Organization (Uganda), Ethiopian Biodiversity Institute (Ethiopia), Southern African Development Community (Zambia), and others accounted for the remaining 56.3% (FAO, 2010). These finger millet collections are rich in rare alleles for target traits from which researchers and crop breeders can obtain the genetic materials to expedite their work in a sustainable way. However, a systematized usage of genebank accessions has not progressed very far in finger millet research and breeding programs due to the scanty information. The ongoing initiative by ICRISAT and other collaborators to sequence several collections to expedite its use for breeding. It is our view that this effort will provide an opportunity to mine novel alleles for breeding next generation of finger millet varieties.

Blast disease caused by *M. oryzae* is the most important disease affecting the growth and yield of finger millet. The fungus infects

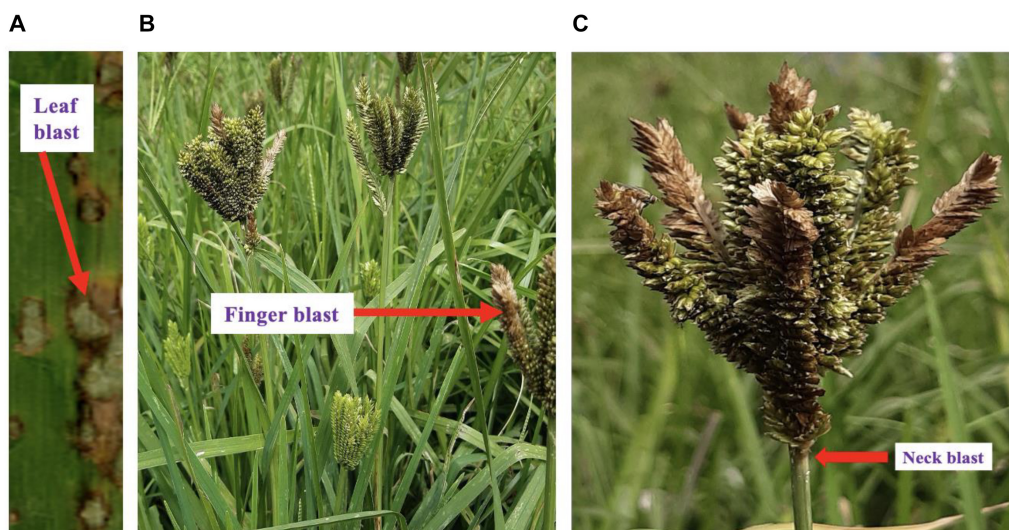


FIGURE 1 | Blast disease symptoms on finger millet (A) blast infection of the finger millet (B), infection on leaves (C), panicle and neck blast symptoms.

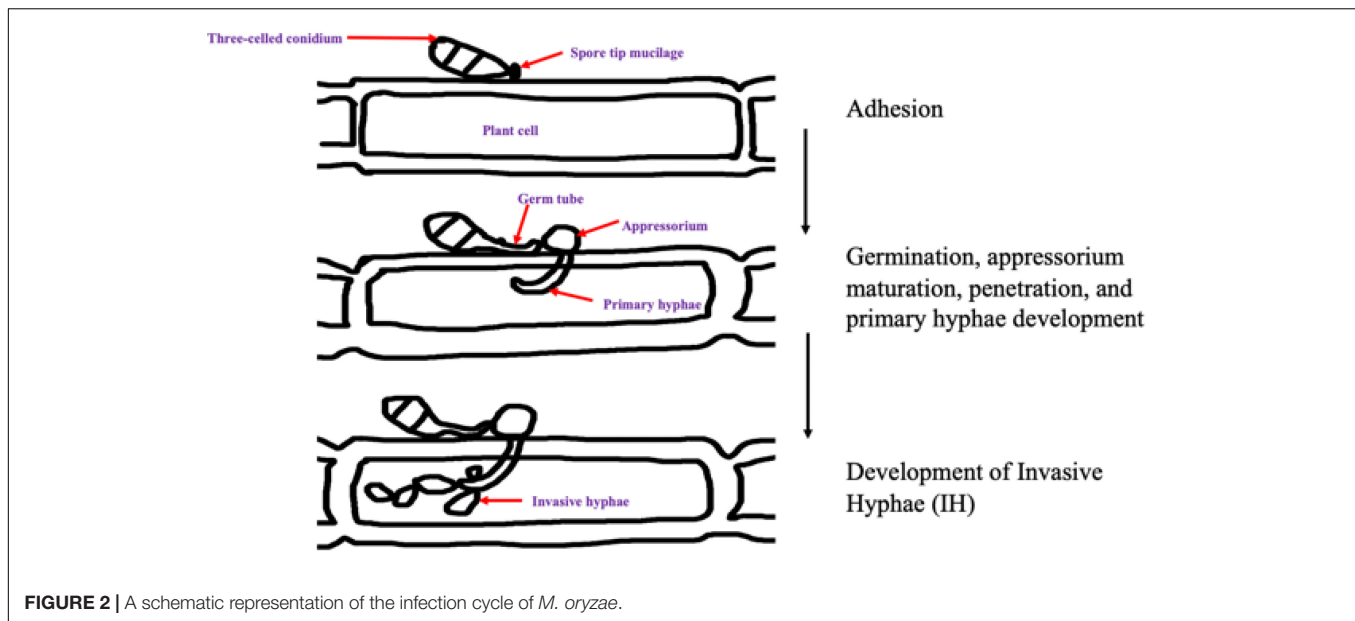
many other economically important crops, such as rice (Mentlak et al., 2012), foxtail millet (Sharma et al., 2014), barley (Tufan et al., 2015), wheat (Cruz and Valent, 2017), and other grass plants within the Poaceae family (Wang et al., 2011; Han et al., 2018). Under favorable conditions, the pathogen infects leaves, stem, collar, node, neck, fingers, and roots, causing substantial crop losses in all finger millet-growing areas. Blast infections occur mostly on the leaves, with the first symptoms of the disease appearing as small gray or brownish dots on the leaves. Following 2–3 days of infection, the dots grow into diamond-shaped lesions of 1.5 cm in length and 0.3–0.5 cm in width, with a white or grayish center (Figure 1). The head blast significantly reduces finger length, seed weight, number of seeds per finger, and total grain yield (Mgonja et al., 2013). The yield losses due to blast have been reported to be between 30 and 50% in large rice-producing areas under favorable environmental conditions (Correa-Victoria and Zeigler, 1993; Skamnioti and Gurr, 2009). Efforts are on-going to develop finger millet varieties with blast resistance. Therefore, continuous studies on blast disease are essential to overcome this disease and thereby sustain finger millet production in the future.

Biology and Pathogenicity of *M. oryzae*

Magnaporthe oryzae is plant-pathogenic filamentous ascomycete fungi that belongs to the *Pyricularia* family. Ascomycete fungi are host specific and causes blast disease on more than 50 cultivated and wild monocot plant species (Gladieux et al., 2018). Based on the polyphyletic nature of *Pyricularia* genus, it is believed that reproduction in the finger millet blast fungus is asexual, producing clonal populations (Figure 2). *M. oryzae* is heterothallic and both mating types *MAT1* and *MAT2* occur in single mating type gene that has two alleles, *MAT1-1* and *MAT1-2*. However, isolates from single agro-ecological region are usually of only one mating type and where the both mating type idiomorphs occur, the strains

cannot interbreed (Takan et al., 2012). Despite this observation, occurrence of highly fertile and hermaphrodite isolates, and isolates harboring the two mating types has been reported (Saleh et al., 2014). Finding suggest that sexual recombination may contribute to genetic variability. Differences have been observed in the two mating types in septoria leaf blotch wheat pathogen *Mycosphaerella graminicola* isolates with *MAT1-1* isolates having significantly greater pathogenicity than *MAT1-2*. On contrary, *M. oryzae* isolates with *MAT1-1* and *MAT1-2* idiomorphs were found to have similar pathogenicity on different monogenetic lines of rice. This has not been tested on finger millet blast strains, and similar results are hypothesized. Despite this suggestion, the finger millet blast fungus sexual recombination, pathogenicity variation, habitat adaptation, and fitness need to be investigated. Further, the phylogenetic evolution and geographic transmission patterns of the finger millet blast pathogen need to be explored.

Magnaporthe oryzae infects the host in two stages, the biotrophic stage where it obtains nutrients from live cells and a necrotrophic stage where it obtains nutrients from dead cells (Park et al., 2009). During infection, conidia attach to the host leaf surface by adhesive secretions released from the apical part of the spore tip during hydration. Subsequently, the spore anchors itself tightly to the hydrophobic (non-stick) finger millet surface to allow germination. After that, the conidia produce germ tubes which form a melanized appressoria. A mature appressorium then breaks the leaf cuticle by creating cellular turgor pressure through the accumulation of compatible solutes such as glycerol and secretion of cell wall degrading enzymes; therefore, gaining entry into the epidermal cells (Marcel et al., 2010; Mentlak et al., 2012). Once inside to the host tissues, *M. oryzae* spreads to adjoining cells through the plasmodesmata without causing any perceptible alteration to the cell walls of the host (Galhano and Talbot, 2011). Under favorable conditions of high humidity, the fungus sporulates abundantly from disease lesions, permitting the



disease to quickly spread to adjacent finger millet plants and its relatives by wind and water droplets (Mentlak et al., 2012).

Management Strategies of Finger Millet Blast Disease

Management of blast disease is a challenging issue, and its control relies on three broad strategies various farming practices, application of chemical and biological agents, and breeding of blast resistant varieties. Some of the cultural and farming practices that have been applied to control blast disease include planting time, spacing, crop rotation, nutrient management (nitrogen and silicon), and water management (Mgonja et al., 2013). These farming practices are commonly used by resource-poor subsistence farmers in developing countries who cannot afford other methods of disease management. Preventive and curative chemicals, such as organophosphorus fungicides, critical elements in effective blast disease management not only in finger millet but also in other crops. Treatment of planting seeds with systemic fungicides such as tricyclazole and application of foliar sprays of edifenphos or kitazin with first at the time of ear emergence or application of carbendazim followed by mancozeb after seven to ten days at the same time been demonstrated to be effective in controlling blast disease (Nagaraja et al., 2007). The indiscriminative nature of chemicals often leads to the development of resistance in phytopathogens. Despite their effectiveness, agrochemicals pose potential risks to human health, food safety and the environment. The growing global concern on the environment, coupled with a strong drive to sustainable agriculture, has led to the advancement of non-chemical alternative strategies methods to control blast disease. Biological control (use of microbial antagonists to suppress diseases) of finger millet blast disease has been considered a viable and sustainable alternative method to synthetic chemical fungicides. Bioinoculants offers multiple beneficial aspects such as the production of quality grains, protection of crops against

biotic and abiotic stresses; soil fertility enhancement and are environmentally safe. Biological agents have successfully been used to manage fungal diseases, such as powdery mildew (Mgonja et al., 2007; Anand et al., 2010), verticillium wilt disease (Yuan et al., 2017), and anthracnose (Hernandez-Montiel et al., 2018).

Biological control (use of microbial antagonists to suppress diseases) of finger millet blast disease has been considered a viable and sustainable alternative method to synthetic chemical fungicides. Bioinoculants offers multiple beneficial aspects such as the production of quality grains, protection of crops against biotic and abiotic stresses; soil fertility enhancement and are environmentally safe. Biological agents have successfully been used to manage fungal diseases such as powdery mildew (Mgonja et al., 2007; Anand et al., 2010), verticillium wilt disease (Yuan et al., 2017), and anthracnose (Hernandez-Montiel et al., 2018). Attempts have been made to control blast disease using bioinoculants. For example, indigenous rhizosphere *Pseudomonas* sp. strain MSSRFD41, had been shown to control blast disease and promote the growth of finger millet *in vitro* but this approach has not been applied at natural field conditions (Sekar et al., 2018). Similarly, endophytic *Bacillus tequilensis* GYLH001 inhibits the growth of *M. oryzae* rice blast and thus has a high potential application as a bioinoculant for control of rice blast pathogen (Li et al., 2018). These examples offer encouraging results, affirming that bioinoculants can significantly contribute to limit the damage caused by blast disease. However, no effective bioinoculant has been formulated and widely adopted for effective biological control of blast disease pathogen. Effective adoption of bioinoculants requires an improvement in understanding of the complex plant-microbe interactions, and an efficient and stable antagonist for the pathogen under different agroecological conditions must be obtained for this control strategy to be realized. More efforts should also be done to authenticate the currently available outcomes setting up effective formulations and application protocols and deepen

the knowledge and awareness on the value of biocontrol agents. Like other blast control strategies, it is our view that biocontrol should not be used alone, but should be implemented in an integrated management framework for sustainable protection of finger millet from *M. oryzae*.

In the current biotechnology era, the breeding of blast-resistant varieties offers the best cost effective and reliable approach for the management of finger millet blast disease, especially in developing countries dominated by subsistence farming. A number of improved finger millet varieties such as IE4795, IE 1055, IE 2821, IE 2872, IE 4121, IE 4491, IE 4570, IE 5066, IE 5091, and IE 5537 with broad-spectrum resistance to blast pathogen coupled with desirable agronomic traits such as early flowering, medium stalk length, and high yields have been developed through conventional breeding methods in eastern Africa and Asia. Such The development and deployment in several production systems of these varieties has been a collaborative work of international research institutions and national finger millet breeding programs. The use of such varieties in integrated management of blast disease is desirable because they require minimal fungicides, subsequently lowering production cost. From our unpublished data on appraisal of occurrence, impact, risk factors and farmers' knowledge and attitudes of finger millet blast disease in Kenya, results showed that many farmers did not know the existence of blast resistant lines and this observation could be the same in all other finger millet growing regions. Farmers should therefore be made aware of the benefits planting resistant lines as opposed to their preferred cultivars. Local finger millet landraces and their wild relatives are commonly used as sources of variation of introgression and hybridization to incorporate the range of useful adaptations for disease resistance into cultivated finger millet. Even though genetic resistance will continue to be the main strategy for control of blast disease, the success of this method is short-lived due to the instability of the *M. oryzae* genome, especially the fast-evolving genes, leading to the breakdown of resistance under field conditions (Mgonja et al., 2013). Abiotic stresses, such as drought, have also displayed a partial or complete breakdown of resistance (Gupta et al., 2017). To overcome this challenge, gene stacking/pyramiding and identification of new partial resistance (*R*) genes against finger millet blast disease are an important goal of finger millet breeding. Staking of multiple *R*-genes or the alleles of a major *R*-gene, which recognize the unique set of *M. oryzae* strains through a conventional breeding approach or transgenesis, has been considered for the attainment of dynamic and durable resistance against different strains of the pathogen (Das et al., 2017; Kumari et al., 2017). Modern biotechnological techniques are simpler, cost-effective, can be performed over a short period, and are more efficient than classical breeding methods.

Breeding Approaches for Improvement of Disease-Resistant Varieties

The current breeding approaches combine two or more objectives which include increasing grain yield, improving resistance to resistance to various biotic and abiotic stresses and enhancing

nutritional quality. Therefore, new finger millet varieties that combine all these traits are desirable. Due to the severity of blast disease, conventional breeding approaches for transferring robust and durable resistance to *M. oryzae* into adapted finger millet germplasm has been a goal of many breeding programs (Upadhyaya et al., 2011; Mgonja et al., 2013). A major challenge for finger millet breeding is that different inheritance models that have been published among the sources of resistance to *M. oryzae* due to the pathogen's specificity. The drawbacks of conventional breeding due to genetic drag and erosion, reproductive hindrances and longer period it takes has necessitated the need for novel breeding methods. Although some achievements have been achieved through conventional breeding strategies, dynamic, efficient, versatile, and contemporary tools and resources must be continually be developed and applied in order to create the necessary paradigm shift needed in finger millet research and breeding.

The rapid advancement in next-generation sequencing techniques together with the declining associated costs and high-performance computation, have resulted to extensive discovery of numerous genomic resources in plants and other organisms. The wealth of information emanating from the post-genomic era has enabled a better understanding of the physiological, biochemical, and molecular mechanisms involved in genotype and its relationship with the phenotype especially for complex traits, facilitated systematic improvement of crop breeding, and allowed for the efficient use of genetic resources. Novel DNA-driven breeding techniques such as marker-assisted selection (Ghatak et al., 2017), gene pyramiding (Chen et al., 2008; Singh et al., 2013), marker-assisted backcross breeding, (Valarmathi et al., 2019; Ponnuswamy et al., 2020), speed breeding technology (Chiurugwi et al., 2019; Hickey et al., 2019) and a combination of them have been utilized in several crops such as rice, soya bean, maize and wheat. To our knowledge, no finger millet variety has been developed and released based on marker-assisted selection (MAS) technique to date, despite the potential of MAS in other cereal crops improvement has been demonstrated for important traits such as bacterial blight resistance in rice (Pandey et al., 2013). Application of these contemporary approaches will ultimately expedite finger millet breeding efforts against blast disease. However, a lot of knowledge is required before full application of molecular breeding in finger millet as most of the available data for the crop currently is on diversity studies and limited QTLs.

Plant genetic engineering which encompasses genetic transformation and genome editing have opened new avenues to modify crops and provided solutions to solve specific needs, establishing it as one of the most important and dynamic biotechnological tools to revolutionize agriculture. This technology can integrate foreign genetic material into different plant cells to produce transgenic plants with new desirable traits, such as drought tolerance, pest and disease resistance, and quality improvement. To circumvent the controversy of genetically modified organisms, innovation in genome editing tools that cause genome changes without producing transgenic plants are currently being explored (Luo et al., 2015; Miroshnichenko et al., 2019).

As an orphaned crop, the status of finger millet genetics and genomics still lags behind that of 204 other food cereal crops, such as maize, rice, wheat, and barley, and even other small millets due to limited research interests and investments. Few biotechnological approaches have been tested on finger millet for crop improvement (Sood et al., 2019). However, this reality is rapidly evolving as the cost of technologies decreases, leading to an exponential decline in the cost of the generation of new knowledge. A vast reservoir of more than 28,041 finger millet germplasm is available in various institutions worldwide for genetic and breeding research (Ceasar et al., 2018). The long-awaited whole-genome sequence and annotation will trigger much higher-resolution research on functional genomics, proteomics, comparative genomics and forward and reverse genetics to unravel the molecular mechanisms mediating major agronomic traits, such as yield, grain quality, abiotic stress tolerance, and pest and disease resistance. Subsequently, the genomic knowledge will be transferred into crop productivity through molecular breeding and better agronomic husbandry.

With the expected large amounts of omics sequencing data, finger millet scientists need to prepare the emerging opportunities and challenges for multi-omics big data integration by means of artificial intelligence for a feasible improvement approach. Studies on genome collinearity show a high genomic synteny between finger millet and rice, foxtail millet and maize, in that order (Hittalmani et al., 2017; Pandian and Ramesh, 2019). Blast resistance in finger millet has been studied using comparative genomics, and different approaches have been used to genetically improve finger millet for effective, durable resistance to important diseases. Several *R*-genes and quantitative trait loci (QTL) in finger millet linked to the blast pathogen have been reported and sequenced (Ramakrishnan et al., 2016; Odeny et al., 2020). These genes and QTL show high sequence similarity in rice and barley, signifying a common evolutionary ancestry for these *R*-genes (Ramakrishnan et al., 2016). Similar methods could be used to identify novel alleles for blast resistance through syntenic studies with the data available from rice, pearl millet, barley, and other related plants. To date, there is no literature reporting mutation breeding in finger millet to generate new varieties. Still, this avenue can be explored given the successful mutation breeding of rice, which has focused mainly on grain quality and taste, agronomic traits, and resistance against pests and diseases. Traditional finger millet landraces have been widely used as genetic resources of breeding programs (Mirza and Marla, 2019). However, the lengthy 10–15-year breeding cycle from crossing to variety release slows the progress.

The majority of the traits, such as resistance to blast disease, are polygenic, and this poses a major challenge when combining large numbers of traits. To supplement conventional breeding, shorten breeding cycles, and accelerate research activities, powerful tools, such as speed breeding protocols, which accelerate plant growth and development, may be explored in finger millet breeding strategies (Watson et al., 2018). Speed breeding technology is able to achieve up to six generations per year of barley (Hickey et al., 2017), wheat (Watson et al., 2018), and oats (González-Barrios et al., 2020), presenting a robust tool to reduce the long period of breeding cycles effectively. Several

speed breeding protocols that utilize prolonged photoperiods and controlled temperatures to accelerate growth and development have been developed for the world's major cereals. Finger millet is a tropical, short-day plant, but its speed breeding protocol is yet to be developed. In order to realize the actual and potential opportunities of speed breeding technology, it is essential to optimize the parameters at a low cost for finger millet. Being time and resource saving, speed breeding technology will accelerate research, improve stability, and increase global finger millet production to meet food security demands of the increasing population.

Blast Disease Improvement Due to Markers and Genotyping Systems

Molecular markers are highly treasured in plant genetics. Over the years, molecular markers have played a prominent and versatile role in finger millet breeding for cultivar improvement, taxonomy, population genetics, plant physiology, and genetic engineering (Veluru et al., 2020). Characterization of finger millet using isozyme makers found fixed heterozygosity at several loci that was identical across the examined accessions (Werth et al., 1994). DNA-based makers such as restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), simple sequence repeat (SSR), expressed-sequenced tag (EST), and markers have been used to generate genetic map of finger millet (Dida et al., 2007; Babu et al., 2014a). Both isozyme and DNA marker analyses have indicated a low variation within cultivated finger millet. Highly variable makers will therefore be required for their application in the crop's breeding.

The advancement in high-throughput sequencing techniques has facilitated sequencing for whole genome sequencing and re-sequencing projects, generating large volumes of sequence data quickly and at a reasonable cost. Next generation sequencing techniques presents new avenues for high-marker density genotyping procedures such as genotyping by sequencing (GBS), which can be used to unravel large numbers of single nucleotide polymorphisms (SNPs) for species identification, diversity analysis, linkage mapping, and genome-wide association studies (GWAS) (Elshire et al., 2011). Because of its high throughput and robustness, GBS can be used to unravel the close variation of cultivated finger millet genotypes. GBS was successfully applied to establish the genetic diversity, population structure and ploidy level among 112 *Vanilla planifolia* accessions and identified 521,732 SNP markers (Hu et al., 2019). Further, genetic diversity of olive germplasm (*Olea europaea* L.) was achieved through GBS technology (Zhu et al., 2019). These results validate the efficacy of genomics-based tools in species genotyping and demonstrate GBS as an effective marker for cultivar genetic diversity analysis in several cultivated crops, providing a vital tool for genomics-assisted plant breeding. Although the genome sequence for finger millet has not been release so far, considerable gains of GBS have been achieved. For example, Tiwari et al. (2020) discovered genes and QTLs governing seed protein content and related traits in finger millet using SNPs discovered via GBS technology. Likewise, GBS was used to identify genomic regions which govern grain nutritional traits in finger millet and generated 169,365 SNPs and

three subpopulations (Puranik et al., 2020). These two examples signify the utility of GBS in genome wide association analysis in mining of novel fundamental genetic information which is essential for marker-assisted breeding against blast disease. The limitations which may arise is that many of the SNPs identified from blast infection are likely to be associated with broad stress or infection responses.

The main challenge posed by the use molecular markers in plant breeding is the high cost of establishing, maintaining of molecular laboratories and inadequate qualified human resource. Moreover, the huge capital requirement for development of markers another major impediment in use of molecular markers in plant breeding programs especially in many developing countries where finger millet is grown. These countries should put more efforts to surmount these challenges. These challenges could be resolved through establishment of specialty regional and continental molecular laboratories could cut costs and bring synergy in research and plant breeding activities.

Resistant Genes and QTLs for Blast Disease

Molecular marker-based breeding approaches have been valuable in the development of blast resistance and in improving important agronomic traits in crops, such as rice and foxtail millet (Tabien et al., 2002). The majority of these traits, for instance, blast resistance, are under quantitative genetic control (Fu, 2015). Nine blast *R*-genes (*Pita*, *Pi9*, *Pi2*, *Piz-t*, *Pi-kh*, *Pi36*, and *Pi37*) belonging to the nucleotide-binding site-leucine-rich repeat (NBS-LRR) family have been cloned in rice by using different strategies (Ameline-Torregrosa et al., 2008). EST sequences of NBS-LRR regions of finger millet have shown homology with *Pi-kh* and *Pi21*, indicating that rice blast *R*-gene orthologs may be playing a crucial role in conferring resistance in finger millet (Kumar et al., 2016a). Genetic mapping and molecular characterization of quantitative traits enable genome-aided breeding in improving the finger millet crop. The common tools used for analyzing the quantitative traits are association mapping and linkage analysis. Association mapping for resistance has been done using genic-SSR markers strongly linked to blast QTL from the finger millet NBS-LRR region in the identification of QTL for finger blast and neck blast resistance. Babu et al. (2014b) identified five significant QTL for finger blast and neck blast. The finger blast QTL were strongly associated with the genic-SSR primer FMBLEST32 and rice SSR RM262 (Babu et al., 2014a). The FMBLEST32 marker was designed from a *Pi5* rice blast gene known for a relatively broader spectrum resistance to *M. oryzae* (Wang et al., 1994). Ramakrishnan et al. (2016) identified two leaf blast resistance QTL strongly associated with markers UGEP101 and UGEP95 by association mapping.

Gene Pyramiding for Blast Resistance

The notable losses of finger millet to blast disease necessitates the development of highly improved and novel strategies to enhance the capacity of various finger millet varieties that can survive attacks caused by the ever-evolving, mutating *M. oryzae* pathogen while also enduring the variable farming and climatic

conditions with a high level of grain quality. Although fungicides are an option for the control of blast disease, they are expensive or not readily available to subsistence and smallholder farmers who are the dominant producers of finger millet in the tropics of sub-Saharan Africa and Asia. Finger millet varieties with resistance to *M. oryzae* fungi offer a cost-effective, easy-to-use, and environmentally-friendly management strategy (Pastor-Corrales et al., 1998). Several sources of blast disease resistance have been identified among primary and secondary gene pools of finger millet (Wang et al., 1994; Babu et al., 2014b; Ramakrishnan et al., 2016; Odeny et al., 2020). However, development of varieties with durable blast disease resistance is difficult. In particular, the broad and dynamic virulence diversity of the blast pathogen has been a limiting factor for host-plant resistance to blast disease because it renders the resistant varieties susceptible within a short period (Odeny et al., 2020).

The advancement of molecular methods in plant breeding has significantly broadened the identification of various *R*-genes in finger millet and other important crops of the Poaceae family. The presence of a set of different *R*-genes in the same plant averts the infection from multiple pathogen races, thereby avoiding fungal evolution by preventing recombination between different fungal races. Previous evidence on rice blast disease has shown that integrating a set of different valuable *R*-genes or QTLs into the same plant would block the infection from several pathogen strains, consequently preventing the fungal evolution through the averting recombination between different fungal races (Yasuda et al., 2015). Gene pyramiding involves stacking of multiple-genes, which results in the simultaneous expression of the various genes in the same plant (Delmotte et al., 2016). Although this approach looks promising in compacting blast and other fungal disease in rice and other cereal crops such as maize (Zhu et al., 2018) and wheat (Cruppe et al., 2020), although, it has not been tried in finger millet. An assessment of the performances of three approaches to control root-knot nematode: cultivar mixtures, crop rotation, and pyramiding of *R*-genes in pepper and lettuce under controlled conditions for more than 3 years (Djian-Caporalino et al., 2014). Results from their work demonstrated that pyramiding of different genes conferring resistance to root-knot nematodes in one genotype was more durable and suppressed the emergence of virulent isolates than pyramiding of different genes conferring resistance to root-knot nematodes in cultivar mixtures and crop rotations. These empirical results, together with theoretical considerations of qualitative and quantitative disease resistance and retrospective analysis, pinpoint that gene pyramiding is the most powerful strategy to provide durable resistance to plant pathogens.

However, gene pyramiding has challenges, including compromised efficacy of stacking genes, if critical assumptions are not adhered to and virulence gene mutations occur, which are often independent of one another. Moreover, the masking of gene expression of resistance, genotype \times environment interactions, the phenotypes, and physiological and biochemical penalties linked with *R*-genes, could eviscerate the agronomic performances. The long period required to obtain a successful variety through gene pyramiding is another major impediment, especially to seed companies. This challenge has been resolved

by the new approaches, such as novel sequencing technologies, marker-assisted selection, genetic engineering, and genomic editing. These new methods have aided in the discovery of new, essential *R*-genes with ease and facilitated their combination into a single variety. The breakdown of pyramided genes has also been recorded in several experiments and explained the theoretical projection (Delmotte et al., 2016; Rana et al., 2019). In turn, this creates probabilities of the emergence of multi-virulent pathogen strains, such as *M. oryzae*. Therefore, it is imperative to strike a balance between an economic impact and effective strategies for controlling the disease.

Allele Mining and Blast Resistance Genes

The advancement made in breeding of superior crops has been achieved by gathering of valuable alleles from vast plant genetic resources from different agro-ecological regions of the world. The wild relatives and landraces of crops still have numerous untapped valuable alleles which could be sustainably exploited for development of superior cultivars which are able withstand environmental variations and still retain the preferred qualities. Introgression of novel alleles from wild relatives into cultivated crop varieties such as stripe rust resistant wheat (Liu et al., 2020), tomato against tomato leaf curl virus, late blight and root knot nematodes (Kumar et al., 2019), development of a rice mega rice variety “Tellahamsa” for bacterial blight and blast resistance (Jamaloddin et al., 2020) as proven that specific alleles and their combinations produce dramatic trait changes when introgression into a suitable genetic background. At the moment, no report exists on gene pyramiding on finger millet. More efforts should therefore be done to unravel more new important alleles to continually enrich the genetic potential of crops. The techniques and prospects of allele mining in the genomic era has been extensively reviewed (Kumari, 2018), so it is not discussed in detail here.

Together with other constraints, blast disease causes a yield loss of as high as 100% in areas infested with the pathogen (Mgonja et al., 2013), purporting a need to understand the molecular mechanism of blast resistance and identify *R*-genes for the blast disease. With the advancement of sequencing technology, enormous sequence and expression data has been deposited into various databases. The use of these novel genomic tools has accelerated discovery and annotation of novel genes and further facilitated the development of allele-specific-markers. Due to the scarcity of genomic resources for genetic improvement of finger millets, comparative genomics will play a critical role in analyzing the most useful and essential agronomic traits, like blast resistance (Kumar et al., 2016a). Comparative analysis of finger millet and rice genomes has demonstrated that most of the chromosomes are highly collinear with 85% synteny (Srinivasachary et al., 2007). Synteny relationship between rice and rice mapped blast *R*-genes through association mapping using NBS-LRR EST sequences, *M. grisea* and *Pi* rice genes of rice. Babu et al. (2014b) found that the finger millet blast and neck blast QTL were linked to rice genes, such as *Pi5*, *Pi21*, *Pi-d(t)*, and NBS-LRR. Therefore, these rice blast *R*-genes (*Pi5*, *Pi21*, *Pi-d(t)*, and NBS-LRR) can be targeted for allele mining in finger millet.

Transgenesis for Blast Resistance

The improvement of finger millet using biotechnological tools has lagged when compared with the research made in other major cereals. Genetic engineering of finger millets is essential to improve the nutritional quality and resistance to abiotic and biotic stresses. Improvement of crops through biotechnological techniques depends largely on successful and efficient plant tissue culture protocols that can be categorized into direct organogenesis, indirect organogenesis, and somatic embryogenesis (Loyola-Vargas and Ochoa-Alejo, 2018). Previous studies on finger millet have identified several inherent challenges associated with *in vitro* regeneration, such as the severe recalcitrant nature, polyploidy, and genotypic dependence, which singly or collectively frustrate the plant tissue culture work and, consecutively, the crop improvement systems through transgenesis (Dosad and Chawla, 2016). Plant regeneration in finger millet using different explants in different genotypes has been reported, such as epicotyl (Patil et al., 2009), shoot apical meristem (Babu et al., 2018; Ngetich et al., 2018), mature seeds (Bayer et al., 2014; Pande et al., 2015), and mature embryos (Satish et al., 2016). These protocols provide an opportunity to improve *in vitro* plant regeneration studies in finger millet, although optimization is required for each genotype. To date, there is no literature on *in vitro* regeneration of finger millet using anther culture, protoplast, and protoplasmic fusion. Attempts should also be made to establish a genotype-independent *in vitro* regeneration system for finger millet.

Various protocols have been used for genetic engineering of finger millet, including biolistic (Gupta et al., 2001), microprojectile bombardment (Latha et al., 2005), and *Agrobacterium tumefaciens*-mediated transformation (Ceasar and Ignacimuthu, 2011; Hema et al., 2014). Among them, *Agrobacterium*-mediated transformation is the most successful and frequently used method to deliver DNA for the production of transgenic finger millet. However, there have been few attempts to use these procedures for developing transgenic finger millet lines resistant to blast disease. Latha et al. (2005) produced transgenic finger millet plants resistant to leaf blast disease by using the biolistic transformation technique to introduce a gene coding for an antimicrobial peptide of prawn. In other work, finger millet plants conferring resistance to leaf blast disease were developed via *Agrobacterium*-mediated genetic transformation of a rice chitinase (*chi11*) gene (Ignacimuthu and Ceasar, 2012). Results from these two studies demonstrated a high level of resistance to leaf blast disease in the transgenic plants compared with the untransformed control plants. Blast disease also affects the neck and the fingers, yet there is no reported work on transgenic finger millet resistant to these crop organs. A transgenic approach for developing finger millet with broad and durable blast resistance necessitates screening many potential antifungal genes and pyramiding of possible genes because the fungus is highly variable and can often overcome the deployed blast-resistant cultivars in a short period when resistance is dependent on one major *R*-gene. Stacking multiple *R*-genes into a single variety through genetic transformation is a promising tool for breeding durable and superior resistance, especially when the *R*-genes emanate from different gene clusters

and different host resistance interactions between the *R*-genes and their effector proteins are provided (Dong and Ronald, 2019). The whole-genome sequence of finger millet, which is anxiously being awaited by finger millet research enthusiasts, is expected to be exploited to facilitate the new genome editing tools, especially, the CRISPR/Cas system. CRISPR (clustered regulatory interspaced short palindromic repeats) and its associated proteins (Cas) guide the complex to cleave complementary DNA. The CRISPR/Cas system is revolutionary and innovative tool for plant genome editing because of its simplicity, a wide range of applications, and is cost-effective (Langner et al., 2018). In order to produce only genome edited plants lacking any foreign DNA inserts. It is therefore vital to establish a protoplast-based regeneration system for finger millet to accomplish the goal of producing blast-resistant finger millet and also propel research and innovation in the crop to the next and higher level comparable to rice, wheat, barley, and maize.

Integrating Disease Resistance Genes With Other Crop Disease Control Strategies

In a permanently dynamic world and society, environmental protection and agricultural sustainability remain the core drivers for food security. Despite the technological advancements made over the past two decades, a real food crisis due to plant diseases has emerged as a significant threat to food security worldwide. Integration of durable disease *R*-genes in the advent of pathogen evolution caused by climate change perturbations and other evolutionary pressures provides sturdy protection against crop diseases. It complements a diversified, integrated management of *M. oryzae* because the simultaneous use of m control 'weapons' guarantees maximum returns. Empirical studies and theoretical models demonstrate that an amalgamation of different selective pressures delays the emergence of virulence (Anderson et al., 2019). For example, the durability of an introgressed *R*-gene(s) targeting a pathogenic fungus, such as *M. oryzae*, could be significantly elevated through the application of fungicides targeting that particular pathogen. Generally, all agricultural practices intended to control a given pathogen should theoretically be integrated to increase their respective effectiveness and durability (Anderson et al., 2019; Hu et al., 2020). In cognizance of smallholder farmers who predominantly cultivate finger millet, such combinations may be constrained by financial, technical knowhow, human health, and environmental factors. To achieve the goal of integrated management of blast disease in finger millet, community engagement and extension services, a healthy partnership between all players in the finger millet value chain and training programs should be emphasized to achieve long-term success.

REFERENCES

- Ameline-Torregrosa, C., Wang, B. B., O'Bleness, M. S., Deshpande, S., Zhu, H., Roe, B., et al. (2008). Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant *Medicago truncatula*. *Plant Physiol.* 146, 5–21. doi: 10.1104/pp.107.104588

Conclusion

Effective disease management strategies are crucial to sustaining the production of high-quality crops, as well as reducing the environmental impacts attributable to pathogens and their management measures. *M. oryzae* is the causative agent of blast, the most damaging disease of finger millet, affecting the finger millet production and causing massive yield loss of up to 80% of finger millet yield per annum globally (Mgonja et al., 2013). Information on the genetic identity of *M. oryzae*, as well as its pathogenesis, is important for the precise development of finger millet varieties with different *R*-genes. The development of resistant varieties with durable resistance through the introgression of new genes into a variety is an effective, economical, environmentally friendly, and sustainable approach to controlling the finger millet blast disease. Various molecular tools are available that facilitate the mining of many genes of interest, such as blast *R*-genes and QTL, and identification of their sources with ease, which presents an opportunity for efficient improvement of finger millet through different breeding techniques. Gene pyramiding, allele mining, speed breeding, genetic engineering, genome editing, and other novel molecular breeding approaches present possibilities to attain durable resistance against the blast disease pathogen in finger millet. Blast disease control by combining disease *R*-genes with other methods could be expected to improve the durability of genetic resistance in improved finger millet cultivars.

AUTHOR CONTRIBUTIONS

WM conceptualized the manuscript and critically edited the manuscript for publication. WM and HM undertook the literature review, analysis, and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

WM was supported by the International Centre for Genetic Engineering and Biotechnology (ICGEB) through grant No. CRP/KEN18-01.

ACKNOWLEDGMENTS

The authors thank Prof. Santie de Villiers of Department of Biochemistry and Biotechnology, Pwani University for providing some of the figures.

- Anand, T., Chandrasekaran, A., Kuttalam, S., Senthilraja, G., and Samiyappan, R. (2010). Integrated control of fruit rot and powdery mildew of chilli using the biocontrol agent *Pseudomonas fluorescens* and a chemical fungicide. *Biol. Control* 52, 1–7. doi: 10.1016/j.biocontrol.2009.09.010
- Anderson, J., Ellsworth, P. C., Faria, J. C., Head, G. P., Owen, M. D., Pilcher, C. D., et al. (2019). Genetically engineered crops: importance of diversified integrated

- pest management for agricultural sustainability. *Front. Bioeng. Biotechnol.* 7:24. doi: 10.3389/fbioe.2019.00024
- Babu, B. K., Agrawal, P. K., Pandey, D., Jaiswal, J. P., and Kumar, A. (2014a). Association mapping of agro-morphological characters among the global collection of finger millet genotypes using genomic SSR markers. *Mol. Biol. Rep.* 41, 5287–5297. doi: 10.1007/s11033-014-3400-6
- Babu, B. K., Dinesh, P., Agrawal, P. K., Sood, S., Chandrashekara, C., Bhatt, J. C., et al. (2014b). Comparative genomics and association mapping approaches for blast resistant genes in finger millet using SSRs. *PLoS One* 9:e99182. doi: 10.1371/journal.pone.0099182
- Babu, G. A., Vinoth, A., and Ravindhran, R. (2018). Direct shoot regeneration and genetic fidelity analysis in finger millet using ISSR markers. *Plant Cell Tissue Organ Culture* 132, 157–164. doi: 10.1007/s11240-017-1319-z
- Bayer, G. Y., Yemets, A. I., and Blume, Y. B. (2014). Obtaining the transgenic lines of finger millet *Eleusine coracana* (L.) with dinitroaniline resistance. *Cytol. Genet.* 48, 139–144. doi: 10.3103/S0095452714030025
- Calanca, P. P. (2017). “Effects of abiotic stress in crop production,” in *Quantification of Climate Variability, Adaptation and Mitigation for Agricultural Sustainability*, eds M. Ahmed and C. Stockle (Cham: Springer), 165–180.
- Ceasar, A. S., Maharajan, S. T., Ajeesh, K. T., Ramakrishnan, T. P., Victor, M., Roch, G., et al. (2018). Finger Millet [*Eleusine coracana* (L.) Gaertn.] improvement: current status and future interventions of whole genome sequence. *Front. Plant Sci.* 9:1054. doi: 10.3389/fpls.2018.01054
- Ceasar, S. A., and Ignacimuthu, S. (2011). *Agrobacterium*-mediated transformation of finger millet (*Eleusine coracana* (L.) Gaertn.) using shoot apex explants. *Plant Cell Rep.* 30, 1759–1770. doi: 10.1007/s00299-011-1084-0
- Chandrasekara, A., and Shahidi, F. (2010). Content of insoluble bound phenolics in millets and their contribution to antioxidant capacity. *J. Agric. Food Chem.* 58, 6706–6714. doi: 10.1021/jf100868b
- Chen, H. Q., Chen, Z. X., Ni, S., Zuo, S. M., Pan, X. B., and Zhu, X. D. (2008). Pyramiding three genes with resistance to blast by marker assisted selection to improve rice blast resistance of Jin23B. *Chin. J. Rice Sci.* 22, 23–27.
- Chiurugwi, T., Kemp, S., Powell, W., and Hickey, L. T. (2019). Speed breeding orphan crops. *Theor. Appl. Genet.* 132, 607–616. doi: 10.1007/s00122-018-3202-7
- Correa-Victoria, F. J., and Zeigler, R. S. (1993). Pathogenic variability in *Pyricularia grisea* at a rice blast hot spot breeding site in eastern Colombia. *Plant Dis.* 77, 1029–1035. doi: 10.1094/PD-77-1029
- Cruppe, G., Cruz, C. D., Peterson, G., Pedley, K., Asif, M., Fritz, A., et al. (2020). Novel sources of wheat head blast resistance in modern breeding lines and wheat wild relatives. *Plant Dis.* 104, 35–43. doi: 10.1094/PDIS-05-19-0985-RE
- Cruz, C. D., and Valent, B. (2017). Wheat blast disease: danger on the move. *Trop. Plant Pathol.* 42, 210–222. doi: 10.1007/s40858-017-0159-z
- Das, G., Patra, J. K., and Baek, K. H. (2017). Insight into MAS: a molecular tool for development of stress resistant and quality of rice through gene stacking. *Front. Plant Sci.* 8:985. doi: 10.3389/fpls.2017.00985
- Delmotte, F., Bourguet, D., Franck, P., Guillemaud, T., Reboud, X., Vacher, C., et al. (2016). Combining selective pressures to enhance the durability of disease resistance genes. *Front. Plant Sci.* 7:1916. doi: 10.3389/fpls.2016.01916
- Dida, M. M., Ramakrishnan, S., Bennetzen, J. L., Gale, M. D., and Devos, K. M. (2007). The genetic map of finger millet, *Eleusine coracana*. *Theor. Appl. Genet.* 114, 321–332. doi: 10.1007/s00122-006-0435-7
- Djian-Caporalino, C., Palloix, A., Fazari, A., Marteu, N., Barbary, A., Abad, P., et al. (2014). Pyramiding, alternating or mixing: comparative performances of deployment strategies of nematode resistance genes to promote plant resistance efficiency and durability. *BMC Plant Biol.* 14:53. doi: 10.1186/1471-2229-14-53
- Dong, O. X., and Ronald, P. C. (2019). Genetic engineering for disease resistance in plants: recent progress and future perspectives. *Plant Physiol.* 180, 26–38. doi: 10.1104/pp.18.01224
- Dosad, S., and Chawla, H. S. (2016). *In vitro* plant regeneration and transformation studies in millets: current status and future prospects. *Indian J. Plant Physiol.* 21, 239–254. doi: 10.1104/pp.18.0122
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379
- FAO (2010). *The Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture*. Rome: Food and Agriculture Organization of the United Nations.
- FAOSTAT (2019). *Food and Agriculture Data*. 2019. Rome: Food and Agriculture Organization. Available online at: <http://www.fao.org/statistics/en/>
- Fu, Y. B. (2015). Understanding crop genetic diversity under modern plant breeding. *Theor. Appl. Genet.* 128, 2131–2142. doi: 10.1007/s00122-015-2585-y
- Galhano, R., and Talbot, N. (2011). The biology of blast: understanding how *Magnaporthe oryzae* invades rice plants. *Fungal Biol. Rev.* 25, 61–67. doi: 10.1016/j.fbr.2011.01.006
- Ghatak, A., Chaturvedi, P., and Weckwerth, W. (2017). Cereal crop proteomics: systemic analysis of crop drought stress responses towards marker-assisted selection breeding. *Front. Plant Sci.* 8:757. doi: 10.3389/fpls.2017.00757
- Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., et al. (2018). Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. *mBio* 9:e01219-17. doi: 10.1128/mBio.01219-17
- González-Barrios, P., Bhatta, M., Halley, M., Sandro, P., and Gutiérrez, L. (2020). Speed breeding and early panicle harvest accelerates oat (*Avena sativa* L.) breeding cycles. *Crop Sci.* doi: 10.1002/csc2.20269
- Gupta, P., Raghuvarshi, S., and Tyagi, A. K. (2001). Assessment of the efficiency of various gene promoters via biolistics in leaf and regenerating seed callus of millets, *Eleusine coracana* and *Echinochloa crusgalli*. *Plant Biotechnol.* 18, 275–282. doi: 10.5511/plantbiotechnology.18.275
- Gupta, S. M., Arora, S., Mirza, N., Pande, A., Lata, C., Puranik, S., et al. (2017). Finger millet: a “certain” crop for an “uncertain” future and a solution to food insecurity and hidden hunger under stressful environments. *Front. Plant Sci.* 8:643. doi: 10.3389/fpls.2017.00643
- Han, Y. J., Zhong, Z. H., Song, L. L., Stefan, O., Zh, and Lu, G. D. (2018). Evolutionary analysis of plant jacalin-related lectins (JRLs) family and expression of rice JRLs in response to *Magnaporthe oryzae*. *J. Integr. Agric.* 17, 1252–1266. doi: 10.1016/S2095-3119(17)61809-4
- Hema, R., Vemanna, R. S., Sreeramulu, S., Reddy, C. P., Senthil-Kumar, M., and Udayakumar, M. (2014). Stable expression of mtID gene imparts multiple stress tolerance in finger millet. *PLoS One* 9:e99110. doi: 10.1371/journal.pone.0099110
- Hernandez-Montiel, L. G., Gutierrez-Perez, E. D., Murillo-Amador, B., Vero, S., Chiquito-Contreras, R. G., and Rincon-Enriquez, G. (2018). Mechanisms employed by *Debaryomyces hansenii* in biological control of anthracnose disease on papaya fruit. *Postharvest Biol. Technol.* 139, 31–37. doi: 10.1016/j.postharvbio.2018.01.015
- Hickey, L. T., Germán, S. E., Pereyra, S. A., Diaz, J. E., Ziem, L. A., Fowler, R. A., et al. (2017). Speed breeding for multiple disease resistance in barley. *Euphytica* 213:64. doi: 10.1007/s10681-016-1803-2
- Hickey, L. T., Hafeez, A. N., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C., Tester, M., et al. (2019). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37, 744–754. doi: 10.1038/s41587-019-0152-9
- Hittalmani, S., Mahesh, H. B., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y. R., et al. (2017). Genome and transcriptome sequence of finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18:465. doi: 10.1186/s12864-017-3850-z
- Hu, Y., Resende, M. F., Bombarely, A., Brym, M., Bassil, E., Chambers, A. H., et al. (2019). Genomics-based diversity analysis of *Vanilla* species using a *Vanilla planifolia* draft genome and Genotyping-By-Sequencing. *Sci. Rep.* 9, 1–16. doi: 10.1038/s41598-019-40144-1
- Hu, Y., Wilson, S., Schwessinger, B., and Rathjen, J. P. (2020). Blurred lines: integrating emerging technologies to advance plant biosecurity. *Curr. Opin. Plant Biol.* 56, 127–134. doi: 10.1016/j.pbi.2020.04.011
- Ignacimuthu, S., and Ceasar, S. A. (2012). Development of transgenic finger millet (*Eleusine coracana* (L.) Gaertn.) resistant to leaf blast disease. *J. Biosci.* 37, 135–147. doi: 10.1007/s12038-011-9178-y
- Jamaluddin, M., Durga-Rani, C. V., Swathi, G., Anuradha, C., Vanisri, S., Rajan, C. P. D., et al. (2020). Marker assisted gene pyramiding (MAGP) for bacterial blight and blast resistance into mega rice variety “Tellahamsa”. *PLoS One* 15:e0234088. doi: 10.1371/journal.pone.0234088
- Kumar, A., Babu, B. K., Yadav, S., and Agrawal, P. K. (2016a). Allele mining for resistance gene analogs (RGAs) in crop plants: a special emphasis on blast resistance in finger millet (*Eleusine coracana* L.). *Indian J. Genet. Plant Breed.* 76, 1–9. doi: 10.5958/0975-6906.2016.00001.8

- Kumar, A., Jindal, S. K., Dhaliwal, M. S., Sharma, A., Kaur, S., and Jain, S. (2019). Gene pyramiding for elite tomato genotypes against ToLCV (*Begomovirus* spp.), late blight (*Phytophthora infestans*) and RKN (*Meloidogyne* spp.) for northern India farmers. *Physiol. Mol. Biol. Plants* 25, 1197–1209. doi: 10.1007/s12298-019-00700-5
- Kumar, A., Metwal, M., Kaur, S., Gupta, A. K., Puranik, S., Singh, S., et al. (2016b). Nutraceutical value of finger millet [*Eleusine coracana* (L.) Gaertn.], and their improvement using omics approaches. *Front. Plant Sci.* 7:934. doi: 10.3389/fpls.2016.00934
- Kumar, A., Tomer, V., Kaur, A., Kumar, V., and Gupta, K. (2018). Millets: a solution to agrarian and nutritional challenges. *Agric. Food Security* 7:31. doi: 10.1186/s40066-018-0183-3
- Kumari, M., Rai, A. K., Devanna, B. N., Singh, P. K., Kapoor, R., Rajashekara, H., et al. (2017). Co-transformation mediated stacking of blast resistance genes *Pi54* and *Pi54rh* in rice provides broad spectrum resistance against *Magnaporthe oryzae*. *Plant Cell Rep.* 36, 747–755. doi: 10.1007/s00299-017-2189-x
- Kumari, R. (2018). Allele Mining for Crop Improvement. *Int. J. Pure Appl. Biosci.* 6, 1456–1465. doi: 10.18782/2320-7051.6073
- Langner, T., Kamoun, S., and Belhaj, K. (2018). CRISPR crops: plant genome editing toward disease resistance. *Annu. Rev. Phytopathol.* 56, 479–512. doi: 10.1146/annurev-phyto-080417-050158
- Latha, A. M., Rao, K. V., and Reddy, V. D. (2005). Production of transgenic plants resistant to leaf blast disease in finger millet (*Eleusine coracana* (L.) Gaertn.). *Plant Sci.* 169, 657–667. doi: 10.1016/j.plantsci.2005.05.009
- Li, H., Guan, Y., Dong, Y., Zhao, L., Rong, S., Chen, W., et al. (2018). Isolation and evaluation of endophytic *Bacillus tequilensis* GYLH001 with potential application for biological control of *Magnaporthe oryzae*. *PLoS One* 13:e0203505. doi: 10.1371/journal.pone.0203505
- Liu, R., Lu, J., Zhou, M., Zheng, S., Liu, Z., Zhang, C., et al. (2020). Developing stripe rust resistant wheat (*Triticum aestivum* L.) lines with gene pyramiding strategy and marker-assisted selection. *Genet. Resour. Crop Evol.* 67, 381–391. doi: 10.1007/s10722-019-00868-5
- Loyola-Vargas, V. M., and Ochoa-Alejo, N. (2018). “An introduction to plant tissue culture: advances and perspectives,” in *Plant Cell Culture Protocols*, eds V. M. Loyola-Vargas and N. Ochoa-Alejo (New York, NY: Humana Press), 3–13.
- Luo, S., Li, J., Stoddard, T. J., Baltes, N. J., Demorest, Z. L., Clasen, B. M., et al. (2015). Non-transgenic plant genome editing using purified sequence-specific nucleases. *Mol. Plant* 8, 1425–1427. doi: 10.1016/j.molp.2015.05.012
- Marcel, S., Sawers, R., Oakeley, E., Angliker, H., and Paszkowski, U. (2010). Tissue-adapted invasion strategies of the rice blast fungus *Magnaporthe oryzae*. *Plant Cell* 22, 3177–3187. doi: 10.1105/tpc.110.078048
- Mentlak, T. A., Kombrink, A., Shinya, T., Ryder, L. S., Otomo, I., Saitoh, H., et al. (2012). Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. *Plant Cell* 24, 322–335. doi: 10.1105/tpc.111.092957
- Mgonja, M. A., Lenné, J. M., Manyasa, E., and Sreenivasaprasad, S. (eds). (2007). *Finger Millet Blast Management in East Africa. Creating Opportunities for Improving Production and Utilization of Finger Millet*. Patancheru: International Crops Research Institute for the Semi-Arid Tropics.
- Mgonja, M., Audi, P., Mgonja, A. P., Manyasa, E. O., and Ojulong, O. (2013). *Integrated Blast and Weed Management and Microdosing in Finger Millet: A HOPE Project Manual for Increasing Finger Millet Productivity*. Project report. Hyderabad: International Crops Research Institute for the Semi-Arid Tropics, Patancheru.
- Miroshnichenko, D. N., Shulga, O. A., Timerbaev, V. R., and Dolgov, S. V. (2019). Generation of non-transgenic genome-edited plants: achievements, challenges and prospects. *Biotechnologia* 35, 3–26. doi: 10.21519/0234-2758-2019-35-1-3-26
- Mirza, N., and Marla, S. S. (2019). “Finger millet (*Eleusine coracana* L. Gaertn.) breeding,” in *Advances in Plant Breeding Strategies: Cereals*, eds J. Al-Khayri, S. Jain, and D. Johnson (Cham: Springer), 83–132.
- Nagaraja, A., Kumar, J., Jain, A. K., Narasimhudu, Y., Raghuchander, T., Kumar Bijendar, et al. (2007). *Compendium of Small Millets Diseases*. Bangalore: AICSMIP, UAS, GKVK, 1–80.
- Ngetich, A., Mweu, C., Ngugi, M., Mukami, A., Ojulong, H., and Mbinda, W. (2018). Efficient plant regeneration protocol for finger millet [*Eleusine coracana* (L.) Gaertn.] via somatic embryogenesis. *Afr. J. Biotechnol.* 17, 660–667. doi: 10.5897/AJB2018.16452
- Odeny, D. A., Niazi, A., Tesfaye, K., Lule, D., Wanyonyi, S., and Kunguni, J. S. (2020). “Genomic designing for climate smart finger millet,” in *Genomic Designing of Climate-smart Cereal Crops*, ed. C. Kole (Cham: Springer), 287–307.
- Pande, A., Dosad, S., Chawla, H. S., and Arora, S. (2015). *In-vitro* organogenesis and plant regeneration from seed-derived callus cultures of finger millet (*Eleusine coracana*). *Braz. J. Bot.* 38, 19–23. doi: 10.1007/s40415-014-0102-1
- Pandey, M. K., Rani, N. S., Sundaram, R. M., Laha, G. S., Madhav, M. S., Rao, K. S., et al. (2013). Improvement of two traditional Basmati rice varieties for bacterial blight resistance and plant stature through morphological and marker-assisted selection. *Mol. Breed.* 31, 239–246. doi: 10.1007/s11032-012-9779-7
- Pandian, S., and Ramesh, M. (2019). Decoding of finger millet genome: a milestone of millet genomics. *Signal Transduction Insights* 8, 1–3. doi: 10.1177/1178643418820541
- Park, J. Y., Jin, J., Lee, Y. W., Kang, S., and Lee, Y. H. (2009). Rice blast fungus (*Magnaporthe oryzae*) infects *Arabidopsis* via a mechanism distinct from that required for the infection of rice. *Plant Physiol.* 149, 474–486. doi: 10.1104/pp.108.129536
- Pastor-Corrales, M. A., Jara, C., and Singh, S. P. (1998). Pathogenic variation in, sources of, and breeding for resistance to *Phaeoisariopsis griseola* causing angular leaf spot in common bean. *Euphytica* 10, 161–171. doi: 10.1023/A:1018350826591
- Patil, S. M., Sawardekar, S., Bhawe, S. G., Sawant, S. S., Jambhale, N. D., and Gokhale, N. B. (2009). Development of somaclones and their genetic diversity analysis through RAPD in finger millet (*Eleusine coracana* L. Gaertn.). *Indian J. Genet. Plant Breed.* 69, 132–139.
- Ponnuswamy, R., Singh, A. K., Raman, M. S., venkata Subbarao, L., and Cn, N. (2020). Conversion of partial restorer Swarna into restorer by transferring fertility restorer *Rf* gene (s) through marker assisted back cross breeding (MABB) in rice. *Sci. Rep.* 10:1101. doi: 10.1038/s41598-020-58019-1
- Puranik, S., Sahu, P. P., Beynon, S., Srivastava, R. K., Sehgal, D., Ojulong, H., et al. (2020). Genome-wide association mapping and comparative genomics identifies genomic regions governing grain nutritional traits in finger millet (*Eleusine coracana* L. Gaertn.). *Plants People Planet* 2, 649–662. doi: 10.1002/ppp3.10120
- Que, Y., Xu, Z., Wang, C., Lv, W., Yue, X., Xu, L., et al. (2019). The putative deubiquitinating enzyme MoUbp4 is required for infection-related morphogenesis and pathogenicity in the rice blast fungus *Magnaporthe oryzae*. *Curr. Genet.* 66, 561–576. doi: 10.1007/s00294-019-01049-8
- Ramakrishnan, M., Ceasa, S. A., Duraipandian, V., Vinod, K. K., Kalpana, K., Al-Dhabi, N. A., et al. (2016). Tracing QTLs for leaf blast resistance and agronomic performance of finger millet (*Eleusine coracana* (L.) Gaertn.) genotypes through association mapping and in silico comparative genomics analyses. *PLoS One* 11:e0159264. doi: 10.1371/journal.pone.0159264
- Rana, M., Sood, A., Hussain, W., Kaldete, R., Sharma, T. R., Gill, R. K., et al. (2019). “Gene pyramiding and multiple character breeding” in *Lentils*, ed. M. Singh (Cambridge, MA: Academic Press), 83–124.
- Rasool, R. S., Padder, B. A., Wani, A. A., Shah, M. D., Masoodi, K. Z., Khan, N. A., et al. (2020). *Thyrostroma carpophilum* insertional mutagenesis: a step towards understanding its pathogenicity mechanism. *J. Microbiol. Methods* 171:105885. doi: 10.1016/j.mimet.2020.105885
- Saleh, D., Milazzo, J., Adreit, H., Fournier, E., and Tharreau, D. (2014). South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae*. *N. Phytol.* 201, 1440–1456. doi: 10.1111/nph.12627
- Satish, L., Rency, A. S., Rathinapriya, P., Ceasar, S. A., Pandian, S., Rameshkumar, R., et al. (2016). Influence of plant growth regulators and spermidine on somatic embryogenesis and plant regeneration in four Indian genotypes of finger millet (*Eleusine coracana* (L.) Gaertn.). *Plant Cell Tissue Organ Culture* 124, 15–31. doi: 10.1007/s11240-015-0870-8
- Sekar, J., Raju, K., Duraisamy, P., and Ramalingam, V. P. (2018). Potential of finger millet indigenous rhizobacterium *Pseudomonas* sp. MSSRFD41 in blast disease management-growth promotion and compatibility with the resident rhizomicrobiome. *Front. Microbiol.* 9:1029. doi: 10.3389/fmicb.2018.01029

- Senthil, R., Shanmugapackiam, S., and Raguchander, T. (2012). Evaluation of biocontrol agents and fungicides for the management of blast disease of finger millet. *J. Mycol. Plant Pathol.* 42, 454–458.
- Sharma, R., Girish, A. G., Upadhyaya, H. D., Humayun, P., Babu, T. K., Rao, V. P., et al. (2014). Identification of blast resistance in a core collection of foxtail millet germplasm. *Plant Dis.* 98, 519–524. doi: 10.1094/PDIS-06-13-0593-RE
- Singh, V. K., Singh, A., Singh, S. P., Ellur, R. K., Singh, D., Krishnan, S. G., et al. (2013). Marker-assisted simultaneous but stepwise backcross breeding for pyramiding blast resistance genes *Piz5* and *Pi54* into an elite Basmati rice restorer line ‘PRR 78’. *Plant Breed.* 132, 486–495. doi: 10.1111/pbr.12077
- Skamnioti, P., and Gurr, S. J. (2009). Against the grain: safeguarding rice from rice blast disease. *Trends Biotechnol.* 27, 141–150. doi: 10.1016/j.tibtech.2008.12.002
- Sood, P., Singh, R. K., and Prasad, M. (2019). Millets genetic engineering: the progress made and prospects for the future. *Plant Cell Tissue Organ Culture* 137, 421–439. doi: 10.1007/s11240-019-01587-6
- Srinivasachary, Dida, M. M., Gale, M. D., and Devos, K. M. (2007). Comparative analyses reveal high levels of conserved colinearity between the finger millet and rice genomes. *Theor. Appl. Genet.* 115, 489–499. doi: 10.1007/s00122-007-0582-5
- Tabien, R., Li, Z., Paterson, A., Marchetti, M., Stansel, J., and Pinson, S. (2002). Mapping QTLs for field resistance to the rice blast pathogen and evaluating their individual and combined utility in improved varieties. *Theor. Appl. Genet.* 105, 313–324. doi: 10.1007/s00122-002-0940-2
- Takan, J. P., Chipili, J., Muthumeenakshi, S., Talbot, N. J., Manyasa, E. O., Bandyopadhyay, R., et al. (2012). *Magnaporthe oryzae* populations adapted to finger millet and rice exhibit distinctive patterns of genetic diversity, sexuality and host interaction. *Mol. Biotechnol.* 50, 145–158. doi: 10.1007/s12033-011-9429-z
- Tiwari, A., Sharma, D., Sood, S., Jaiswal, J. P., Pachauri, S. P., Ramteke, P. W., et al. (2020). Genome-wide association mapping for seed protein content in finger millet (*Eleusine coracana*) global collection through genotyping by sequencing. *J. Cereal Sci.* 91:102888. doi: 10.1016/j.jcs.2019.102888
- Tufan, F., Uçarlı, C., and Gürel, F. (2015). Analysis of expressed sequence tags from cDNA library of *Fusarium culmorum* infected barley (*Hordeum vulgare* L.) roots. *Bioinformation* 11:34. doi: 10.6026/97320630011034
- Upadhyaya, H. D., Gowd, C. L., and Reddy, V. G. (2007). Morphological diversity in finger millet germplasm introduced from Southern and Eastern Africa. *J. SAT Agric. Res.* 3, 1–3.
- Upadhyaya, H. D., Ramesh, S., Sharma, S., Singh, S. K., Varshney, S. K., Sarma, N. D., et al. (2011). Genetic diversity for grain nutrients contents in a core collection of finger millet (*Eleusine coracana* (L.) Gaertn.) germplasm. *Field Crops Res.* 121, 42–52. doi: 10.1016/j.fcr.2010.11.017
- Valarmathi, M., Sasikala, R., Rahman, H., Jagadeeshselvam, N., Kambale, R., and Raveendran, M. (2019). Development of salinity tolerant version of a popular rice variety improved white ponni through marker assisted back cross breeding. *Plant Physiol. Rep.* 24, 262–271. doi: 10.1007/s40502-019-0440-x
- Veluru, A., Bhat, K. V., Raju, D. V. S., Prasad, K. V., Tolety, J., Bharadwaj, C., et al. (2020). Characterization of Indian bred rose cultivars using morphological and molecular markers for conservation and sustainable management. *Physiol. Mol. Biol. Plants* 26, 95–106. doi: 10.1007/s12298-019-00735-8
- Wang, G. L., Mackill, D. J., Bonman, J. M., McCouch, S. R., Champoux, M. C., and Nelson, R. J. (1994). RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. *Genetics* 136, 1421–1434.
- Wang, X., Wang, J., Jiang, H., Wang, Y., Zhang, Z., Chai, R., et al. (2011). Pathogenicity of rice blast fungus *Magnaporthe oryzae* on *Brachypodium distachyon*. *Rice Sci.* 19, 252–258. doi: 10.1016/S1672-6308(12)60048-1
- Watson, A., Ghosh, S., Williams, M., Cuddy, W. S., Simmonds, J., Rey, M. D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nature Plants* 4, 23–29. doi: 10.1038/s41477-017-0083-8
- Werth, C. R., Hilu, K. W., and Langner, C. A. (1994). Isozymes of *Eleusine* (Gramineae) and the origin of finger millet. *Am. J. Bot.* 81, 1186–1197.
- Wu, W., Wang, L., Zhang, S., Li, Z., Zhang, Y., Lin, F., et al. (2014). Stepwise arms race between AvrPik and Pik alleles in the rice blast pathosystem. *Mol. Plant Microbe Interact.* 27, 759–769. doi: 10.1094/MPMI-02-14-0046-R
- Yasuda, N., Mitsunaga, T., Hayashi, K., Koizumi, S., and Fujita, Y. (2015). Effects of pyramiding quantitative resistance genes *Pi21*, *Pi34*, and *Pi35* on rice leaf blast disease. *Plant Dis.* 99, 904–909. doi: 10.1094/PDIS-02-14-0214-RE
- Yoshida, K., Saunders, D. G. O., Mitsuoka, C., Natsume, S., Kosugi, S., Saitoh, H., et al. (2016). Host specialization of the blast fungus *Magnaporthe oryzae* is associated with dynamic gain and loss of genes linked to transposable elements. *BMC Genomics* 17:1–18. doi: 10.1186/s12864-016-2690-6
- Yuan, Y., Feng, H., Wang, L., Li, Z., Shi, Y., Zhao, L., et al. (2017). Potential of endophytic fungi isolated from cotton roots for biological control against verticillium wilt disease. *PLoS One* 12:e0170557. doi: 10.1371/journal.pone.0170557
- Zhu, S., Niu, E., Shi, A., and Mou, B. (2019). Genetic diversity analysis of olive germplasm (*Olea europaea* L.) with genotyping-by-sequencing technology. *Front. Genet.* 10:755. doi: 10.3389/fgene.2019.00755
- Zhu, X., Zhao, J., Abbas, H. M. K., Liu, Y., Cheng, M., Huang, J., et al. (2018). Pyramiding of nine transgenes in maize generates high-level resistance against necrotrophic maize pathogens. *Theor. Appl. Genet.* 131, 2145–2156. doi: 10.1007/s00122-018-3143-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Mbinda and Masaki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Physiological and Proteomic Signatures Reveal Mechanisms of Superior Drought Resilience in Pearl Millet Compared to Wheat

OPEN ACCESS

Edited by:

María Serrano,
Miguel Hernández University of Elche,
Spain

Reviewed by:

Mehanathan Muthamilarasan,
University of Hyderabad, India
Xiuli Hu,
Henan Agricultural University, China
Pavel Vítámvás,
Crop Research Institute, Czechia

*Correspondence:

Palak Chaturvedi
palak.chaturvedi@univie.ac.at
Wolfram Weckwerth
wolfram.weckwerth@univie.ac.at

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Crop and Product Physiology,
a section of the journal
Frontiers in Plant Science

Received: 29 August 2020

Accepted: 17 November 2020

Published: 13 January 2021

Citation:

Ghatak A, Chaturvedi P,
Bachmann G, Valledor L, Ramšak Ž,
Bazargani MM, Bajaj P,
Jegadeesan S, Li W, Sun X,
Gruden K, Varshney RK and
Weckwerth W (2021) Physiological
and Proteomic Signatures Reveal
Mechanisms of Superior Drought
Resilience in Pearl Millet Compared to
Wheat. *Front. Plant Sci.* 11:600278.
doi: 10.3389/fpls.2020.600278

Arindam Ghatak^{1†}, Palak Chaturvedi^{1*†}, Gert Bachmann¹, Luis Valledor², Živa Ramšak³,
Mitra Mohammadi Bazargani⁴, Prasad Bajaj⁵, Sridharan Jegadeesan⁶, Weimin Li¹,
Xiaoliang Sun⁷, Kristina Gruden³, Rajeev K. Varshney⁵ and Wolfram Weckwerth^{1,7*}

¹ Molecular Systems Biology Lab (MOSYS), Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria, ² Plant Physiology Lab, Organisms and Systems Biology, Faculty of Biology, University of Oviedo, Oviedo, Spain, ³ Department of Systems Biology and Biotechnology, National Institute of Biology, Ljubljana, Slovenia, ⁴ Agriculture Institute, Iranian Research Organization for Science and Technology, Tehran, Iran, ⁵ Center of Excellence in Genomics & Systems Biology, International Crops Research Institute for the Semi-Arid Tropics, Hyderabad, India, ⁶ Premas Life Sciences, Bengaluru, India, ⁷ Vienna Metabolomics Center (VIME), University of Vienna, Vienna, Austria

Presently, pearl millet and wheat are belonging to highly important cereal crops. Pearl millet, however, is an under-utilized crop, despite its superior resilience to drought and heat stress in contrast to wheat. To investigate this in more detail, we performed comparative physiological screening and large scale proteomics of drought stress responses in drought-tolerant and susceptible genotypes of pearl millet and wheat. These chosen genotypes are widely used in breeding and farming practices. The physiological responses demonstrated large differences in the regulation of root morphology and photosynthetic machinery, revealing a stay-green phenotype in pearl millet. Subsequent tissue-specific proteome analysis of leaves, roots and seeds led to the identification of 12,558 proteins in pearl millet and wheat under well-watered and stress conditions. To allow for this comparative proteome analysis and to provide a platform for future functional proteomics studies we performed a systematic phylogenetic analysis of all orthologues in pearl millet, wheat, foxtail millet, sorghum, barley, brachypodium, rice, maize, Arabidopsis, and soybean. In summary, we define (i) a stay-green proteome signature in the drought-tolerant pearl millet phenotype and (ii) differential senescence proteome signatures in contrasting wheat phenotypes not capable of coping with similar drought stress. These different responses have a significant effect on yield and grain filling processes reflected by the harvest index. Proteome signatures related to root morphology and seed yield demonstrated the unexpected intra- and interspecies-specific biochemical plasticity for stress adaptation

for both pearl millet and wheat genotypes. These quantitative reference data provide tissue- and phenotype-specific marker proteins of stress defense mechanisms which are not predictable from the genome sequence itself and have potential value for marker-assisted breeding beyond genome assisted breeding.

Keywords: climate resilience, senescence, cereals, drought stress, proteomics, stay-green trait, secure food production, marker assisted breeding

INTRODUCTION

Feeding nine billion people with balanced nutritional diet under unpredictable severe weather events is a challenging task. Emerging evidence suggests that the climate change crisis will cause shifts in food production and yield loss, causing a severe threat to food security (Lunt et al., 2016). A key strategy to adapt in a changing climate is to develop elite germplasms that can survive under hostile weather conditions with stable yields and to promote underutilized crop species. Focusing and exploiting the large reservoir of minor and underutilized crop plants would provide a more diversified agricultural system and an alternative healthy food resource, ensuring food, and nutritional security (Mabhaudhi et al., 2019).

The world today relies on a small number of crop species for food, majorly C_3 cereals (wheat, rice, barley), and few C_4 cereals like pearl millet, maize and sorghum (Alexandratos and Bruinsma, 2012). By far, the abundance of genetic resources and potentially beneficial traits of C_4 cereals are neglected. However, to alleviate the food crisis, efforts are ongoing to engineer C_4 traits into C_3 crop species, which can massively increase C_3 crop yields (von Caemmerer et al., 2012; Weissmann and Brutnell, 2012; Wang et al., 2014; Rangan et al., 2016). However, such efforts require an improved understanding of the physiological traits (such as deep rooting, stay-green, and senescence etc.) and system-level analysis to identify the regulatory networks underlying these physiological traits under abiotic stress condition in a comparative manner.

Drought can adversely and drastically affect the agricultural sector. It causes land degradation and biodiversity loss. Every year, around 8.5 million ha of rain-fed land and 1.5 million ha of irrigated lands are affected because of salinization (Hanjra and Qureshi, 2010). Drought induces profound changes at the morphological, physiological and biochemical level in all plant tissues (Anjum et al., 2011), mostly disturbing the complicated relationship between sink and source of plant organs. Upon perception of drought stress a complex response is initiated which includes massive transcriptional reprogramming along with anatomical and physiological alterations which include deep root system, changes in leaf morphology, closure of stomata, cuticular wax thickening, hormone induction, reactive oxygen species (ROS) scavenging, osmolyte synthesis, nitrogen assimilation, and amino acid metabolism (Lamalakshmi Devi et al., 2017). These active processes involve genes, proteins and small molecules (metabolites), which play a crucial role in shaping the final phenotype of the plants (Ghatak et al., 2018; Weckwerth et al., 2020). However, response to drought is species-specific and often genotype-specific (Campos et al., 2004). It also largely depends

upon the duration and severity of water loss, age, and stage of the plant development (Pinheiro and Chaves, 2011).

Proteomics has become a powerful tool for analyzing plant response to various environmental stimuli (biotic and abiotic), especially in the comparative studies of genetically diverse germplasms subjected to drought stress, providing fundamental insights into plant responses to pre-determined stress and biochemical pathways that participate in the acclimatization process. Proper evaluation of the data can contribute to the identification of the potential candidates, which are then correlated with the quantitative trait loci (QTLs). These candidates can be further integrated into the marker-assisted breeding strategy to enhance the selection of plants with desired traits (Tuberosa and Salvi, 2006). Several studies were performed to understand the effect of drought stress on crop plants at proteomics level (Riccardi et al., 2004; Ford et al., 2011; Komatsu et al., 2014; Chmielewska et al., 2016; Ghatak et al., 2016; Ghatak et al., 2017a,b; Michaletti et al., 2018; Rodziejewicz et al., 2019).

Wheat (*Triticum aestivum* L.) and pearl millet (*Pennisetum glaucum* (L.) R. Br.) are most important cereal crops. Wheat is a C_3 cereal crop with a hexaploid genome (~ 17 Gb) (Appels et al., 2018). It is a food source of > 50% world population. The yield of wheat is severely compromised under harsh climatic condition, especially drought (Ahmed et al., 2020). Contrastingly, pearl millet is a C_4 grass highly cross-pollinated diploid ($2n = 2x = 14$) with a relative genome of 1.79 Gb and high photosynthetic efficiency (Varshney et al., 2017). It is an underutilized crop, despite its immense nutritional potential which has not been tapped. Unlike wheat, pearl millet is cultivatable in areas with drought, low soil fertility, high salinity, low pH or high temperature. As compared to other cereals, pearl millet showed greater ceiling temperatures for grain yield, making it a climate-resilient crop suitable for semi-arid regions of the world (Varshney et al., 2017). Projected changes in crop yields owing to climate change demand a paradigm shift to enhance the cultivation and distribution of such crops in the market reducing the burden of the crops with high commercial value, e.g., wheat and maize without compromising their nutritional importance. However, there is a lack of studies that provide insights into the molecular machinery underlying stress tolerance in millets in comparison to other important cereals.

To address this aim, in the present study comparative physiological and proteome changes were evaluated in the leaves, roots and seeds of two different pearl millet and wheat genotypes from different geographical origin subjected to drought: (1) to identify physiological traits associated with tolerance to the deleterious effect of drought stress, (2) to

characterize physiological traits such as photosynthetic activity, root length, seed weight and weight of the plant with contrasting degree of drought tolerance, (3) to explore the implications of drought stress on proteome, and identify tissue-specific (roots, leaves and seeds) drought stress-responsive proteins which attribute to the stress tolerance of these crops, (4) to identify and compare abundance profile of the proteins involved in C₄ pearl millet photosynthetic metabolism and wax biosynthesis with the orthologous proteins present in C₃ wheat, and (5) to integrate physiological and biochemical parameters (i.e., identified proteins) using multivariate analysis to obtain a comprehensive picture of the plants “physiological trait/proteome levels” under drought stress.

MATERIALS AND METHODS

Plant Material, Growth Conditions, and Drought Treatment

Two different genotypes of pearl millet (843-22B and ICTP8203) and spring wheat (White Fife and TRI 5630) from different geographic origins (Table 1A) were selected for this study: wheat genotypes, one originating from the United Kingdom, accession number TRI 5357 (White Fife, here indicated as UK), the other from Iran, accession number TRI 5630 (indicated as IR). Pearl millet genotypes originated from different states of India primarily used for breeding and research. Seeds were obtained from the gene bank repository of International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India and IPK, Germany. The experiment started in February and concluded in July. Plants were grown in a controlled condition: the temperature was max 30°C during the daytime, 26°C at nighttime ($\pm 2^\circ\text{C}$). Relative air humidity was 60% during the day, and 80% at nighttime; the light was provided by metal halide lamps (HRI-TS 250W/NDL Neutral white, Radium, Germany) at an intensity of 220 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ (7 a.m. to 9 p.m.). The plants were grown in custom made cylindrical pipes (Ghatak et al., 2016). Each pipe was made of 5 polyethylene segments (15 cm each) amounting to a total height of 75 cm with an inner diameter of 10.3 cm. The total soil volume was 6.25 L. Each tube had two access openings (one in the upmost segment, and one in the 2nd segment from the bottom) for monitoring soil dehydration (monitored by measuring the volumetric soil water content (volume of water/total volume ratio), in percentage) and soil temperature by means of 1% accurate theta probes by (ADC ML3TM) sensors. Soil mixture consisted of three parts of potting ground (peat, humus), 2 parts of sand, 1 part of styromull (Royal Brinkman, the

Netherlands) and 0.1 % NPK was added as initial fertilizer and no pesticides were used.

The irrigation was adapted to the plant physiological needs, i.e., shoot/root development and evapotranspiration, being higher for wheat, and lower for pearl millet (Figure 1A). As a consequence, the wheat plants were kept at $\sim 32\%$ of soil volume (71.11% of field capacity), and pearl millet plants at $\sim 24\%$ of soil volume (53.33% of field capacity). The drought stress began when the plants reached the developmental stage of phase 51–53 on the BBCH scale, which was achieved in 8 weeks for pearl millet and 10 weeks for wheat. The difference in soil water content between control and stressed plants was the first indication of the drought imposed. The plant material (roots, leaves and seeds) were collected considering four biological replicates in each condition (control and stress) for proteomic analysis after 13-days of drought period. The harvested samples were frozen in liquid nitrogen to stop any enzymatic activity. The tissue samples were ground in liquid nitrogen using mortar and pestle. Pulverized tissues were stored at -80°C until further analysis. Table 1B provides the details of the genotypes, harvested tissues and their abbreviations used in the manuscript text, figures and tables.

Physiological Measurements

Stomatal Conductance, Leaf Chlorophyll Content, and Chlorophyll Fluorescence

The effect of drought stress was examined by measuring stomatal conductance (g_s) ($\text{mmol m}^{-2} \text{s}^{-1}$) using PWMR-4 porometer (PP Systems, United States), leaf chlorophyll content using SPAD chlorophyll meter (SPAD 502, Minolta, Tokyo), and chlorophyll fluorescence (F_v/F_m) using plant efficiency analyzer (PEA) (Handy PEA, Hansatech Instruments, King's Lynn, United Kingdom). Mature and fully expanded green leaves were used for the measurement at regular interval until the drought treatment was completed. The measurements were performed non-destructively on plant attached leaves.

Plant Weight (Biomass), Root Length, Panicle/Spike Characteristic (Numbers of Panicle/Spike per Biological Replicate), Seed Weight, and Harvest Index

For the measurement of plant weight (biomass), the cylindrical pipes were dismantled, and the intact plant was removed carefully from the soil. Panicles and spikelets on culms were counted; seed weight was recorded before putting them into liquid nitrogen. Harvest index (HI) was calculated according to Schauer et al. (2006) with the formula:

Harvest index (HI) % = [Total yield (i.e., seed weight)/Total yield + plant weight (i.e., biomass)] $\times 100$

Protein Extraction and Pre-fractionation

The total protein from roots, leaves and seeds was extracted according to Chaturvedi et al. (2013) and Valledor and Weckwerth (2014). In brief, homogenized tissue weighed (20 mg for root, leaf and seed tissue, respectively), and suspended in 200 μL of protein extraction buffer [100 mM Tris- HCl, pH 8.0; 5% SDS, 10% glycerol; 10 mM DTT; 1% plant protease inhibitor

TABLE 1A | Description of the genotypes used in the study.

Wheat <i>Triticum aestivum</i> L.		Pearl millet <i>Pennisetum glaucum</i> (L.) R. Br.	
United Kingdom (UK)	IRAN (IR)	Telangana, India	Maharashtra, India
White Fife	TRI 5630	843-22B	ICTP8203
Sensitive	Tolerant	Sensitive	Tolerant

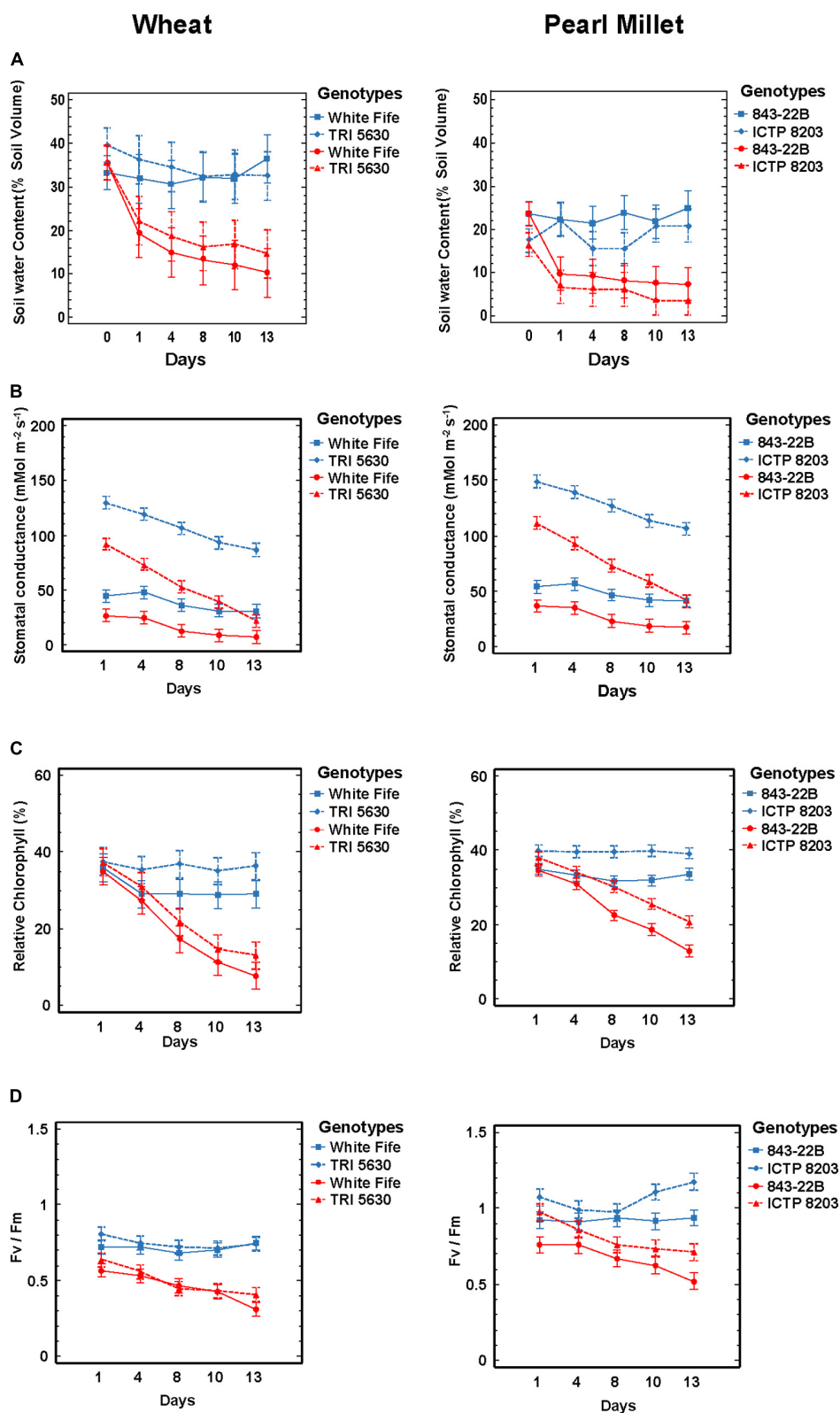


FIGURE 1 | Physiological parameters. **(A)** Soil moisture content was measured using sensors ML 3 ThetaProbe. **(B)** Stomatal conductance was measured using PWMR-4 Porometer. **(C)** Leaf chlorophyll content was determined using SPAD chlorophyll meter. **(D)** Chlorophyll fluorescence (F_v/F_m) was determined using plant efficiency analyzer (PEA). All the analysis was performed under control and stress condition in pearl millet and wheat genotypes (color indication: blue—control and red—drought stress; lines: dotted lines—tolerant genotypes and direct line—sensitive genotypes).

TABLE 1B | Details of the genotypes, tissues harvested and abbreviations used in the figures and tables.

Cereal crops	Genotype name	Genotype	Conditions applied	Tissues harvested	Abbreviations for figures and tables
Wheat	White Fife (indicated as UK)	Sensitive	Control (C)	Root (R)	UK-C-R
				Seed (S)	UK-C-S
			Stress (St)	Leaf (L)	UK-C-L
				Root (R)	UK-St-R
				Seed (S)	UK-St-S
	TRI 5630 (indicated as IR)	Tolerant	Control (C)	Leaf (L)	UK-St-L
				Root (R)	IR-C-R
			Stress (St)	Seed (S)	IR-C-S
				Leaf (L)	IR-C-L
				Root (R)	IR-St-R
Pearl Millet (PM)	843-22B	Sensitive (S)	Control (C)	Seed (S)	IR-St-S
				Leaf (L)	IR-St-L
			Stress (St)	Root (R)	PM-S-C-R
				Seed (S)	PM-S-C-S
				Leaf (L)	PM-S-C-L
	ICTP8203	Tolerant (T)	Control (C)	Root (R)	PM-S-St-R
				Seed (S)	PM-S-St-S
			Stress (St)	Leaf (L)	PM-S-St-L
				Root (R)	PM-T-C-R
				Seed (S)	PM-T-C-S
			Stress (St)	Leaf (L)	PM-T-C-L
				Root (R)	PM-T-St-R
				Seed (S)	PM-T-St-S
				Leaf (L)	PM-T-St-L

cocktail (Sigma P9599)] and incubated at room temperature for 5 min followed by incubation for 2.5 min at 95°C and centrifugation at 21,000 × g for 5 min at room temperature. The supernatant was carefully transferred to a new tube. Two-hundred microliters of 1.4 M sucrose were added to the supernatant and proteins were extracted twice with 200 µL TE buffer-equilibrated phenol followed by counter extraction with 400 µL of 0.7 M sucrose. Phenol phases were combined and subsequently mixed with 2.5 volumes of 0.1 M ammonium acetate in methanol for precipitation of proteins. After 16 h of incubation at −20°C, samples were centrifuged for 5 min at 5,000 × g. The pellet was washed twice with 0.1 M ammonium acetate, once with acetone and air-dried at room temperature. The pellet was re-dissolved in 6 M Urea and 5% SDS, and protein concentration were determined using the bicinchoninic acid assay (BCA method). Proteins were pre-fractionated by SDS-PAGE. Forty micrograms of total protein were loaded onto the gel. Gels were fixed and stained with methanol: acetic acid: water: Coomassie Brilliant Blue R-250 (40:10:50:0.001). Gels were destained in methanol: water (40:60).

Protein Digestion and LC–MS/MS

Gel pieces were destained, equilibrated and digested with trypsin, desalted and concentrated (Chaturvedi et al., 2013). Prior to

mass spectrometric measurement, the tryptic peptide pellets were dissolved in 4% (v/v) acetonitrile, 0.1% (v/v) formic acid. One µg of the digested peptide from each tissue sample (4 biological replicates for each condition) was loaded on a C18 reverse-phase column (Thermo scientific, EASY-Spray 500 mm, 2 µm particle size). Separation was achieved with a 90 min gradient from 98% solution A (0.1% formic acid) and 2% solution B (90% ACN and 0.1% formic acid) at 0 min to 40% solution B (90% ACN and 0.1% formic acid) at 90 min with a flow rate of 300 nL min^{−1}. nESI-MS/MS measurements were performed on Orbitrap Elite (Thermo Fisher Scientific, Bremen, Germany) with the following settings: Full scan range 350–1,800 m/z resolution 120,000 max. 20 MS2 scans (activation type CID), repeat count 1, repeat duration 30 s, exclusion list size 500, exclusion duration 30 s, charge state screening enabled with the rejection of unassigned and +1 charge states, minimum signal threshold 500.

Peptide and Protein Identification

Raw data were searched with the SEQUEST algorithm present in Proteome Discoverer version 1.3 (Thermo, Germany) as described in Valledor and Weckwerth (2014). We have used the following settings in Proteome Discoverer for data analysis which include: Peptide confidence: High, which is equivalent to 1% false discovery rate (FDR), and Xcorr of 2, 3, 4, 5, 6 for peptides of charge 2, 3, 4, 5, 6. The variable modifications were set to acetylation of N-terminus and oxidation of methionine, with a mass tolerance of 10 ppm for parent ion and 0.8 Da for the fragment ion. The number of missed and non-specific cleavages permitted was 2. There were no fixed modifications, as dynamic modifications were used.

For identification, newly annotated pearl millet genome database containing 38,579 genes (Varshney et al., 2017) and UniProt database containing the annotations of 136,865 genes for wheat was used. Peptides were matched against these databases plus decoys, considering a significant hit when the peptide confidence was high. All the MS/MS spectra of the identified proteins and their meta-information were further uploaded to PRIDE repository. Sample codes for the raw files deposited in the PRIDE are provided in **Supplementary Table S1**. Submission details are as follows; Project name: Comparative physiological and proteomic signatures reveal contrasting stay-green and senescence phenotypes in drought tolerant and susceptible pearl millet and wheat genotypes. Project accession: PXD021446.

The identified proteins were quantitated based on total ion count, followed by an NSAF normalization strategy (Paoletti et al., 2006):

$$(NSAF)_k = (PSM/L)_k / \sum_{i=1}^N (PSM/L)_i$$

In which the total number spectra counts for the matching peptides from protein k (PSM) was divided by the protein length (L), then divided by the sum of PSM/L for all N proteins.

Statistical Analysis and Data Integration

Statistical analysis for physiological data points was performed using Statgraphics (ver. 17.2.05) and SIMCA (version 13) for OPLS-DA analysis. For both PCA and OPLS-DA, data were centered and scaled using z-transformation.

Bioinformatics for Functional Annotation

To assign functional descriptions to pearl millet and wheat sequences, BLAST search (Altschul et al., 1997) was performed against Arabidopsis proteins release Araport11 (Cheng et al., 2017), rice (Kawahara et al., 2013), tomato SL3.0_ITAG3.2 (Sato et al., 2012), potato (Xu et al., 2011), and plants in Swiss-Prot (Bateman et al., 2017) using default settings. Every accession was assigned one best match (alignment coverage of shorter sequence $\geq 70\%$; E-value $\leq 10^{-10}$), prioritizing Arabidopsis and rice matches over the rest, when available. Assignment of pearl millet and wheat accession with MapMan plant functional ontology terms (Thimm et al., 2004) was also based on BLAST results against the same databases (alignment coverage of shorter sequence $\geq 70\%$, E-value $\leq 10^{-20}$, bit score ≥ 50 , positives % ≥ 60). Pearl millet and wheat accessions then inherited the BIN assignment from the corresponding best match; unmatched sequenced were assigned BIN 35.2 (not assigned. unknown). To enable visualization of high-throughput experimental results, MapMan and GSEA mapping files were created for each species (available from www.gomapman.org/export/current/, Ramsak et al., 2014).

Gene Family and Phylogenetic Analysis

For gene family analysis of 11 plant species, DIAMOND (Buchfink et al., 2015) was used with an *e*-value cutoff of $\leq 1.0\text{e-}05$. In addition to pearl millet, wheat sequences were downloaded from Swiss-Prot, while PLAZA v4 resource (Van Bel et al., 2018) was used for *Arabidopsis thaliana*, *Brachypodium distachyon*, *Glycine max*, *Hordeum vulgare*, *Oryza sativa* ssp. *japonica*, *Oryza sativa* ssp. *indica*, *Sorghum bicolor*, *Setaria italic*, and *Zea mays*. To reduce redundancy present in wheat sequences, these were pre-processed using CD-HIT (Fu et al., 2012) ($\geq 80\%$ identity; $\geq 80\%$ coverage for shorter and longer sequence compared). The proteins were clustered using OrthoMCL v2.0.9 (Li et al., 2003), to define gene families with paralogs and orthologs. Single copy genes in an OrthoMCL cluster for all species were used to construct a phylogenetic tree in SeaView with muscle for multiple sequence alignment, Gblock to select the conserved regions and PhyML to construct the phylogenetic tree (bootstrap 1000).

Statistics for Proteome Data Analysis

Data were normalized using normalized spectral abundance factor (NSAF) approach and subjected to multivariate (Principal components analysis (PCA), K-means clustering) analysis which was performed using the statistical toolbox COVAIN in MATLAB (Sun and Weckwerth, 2012) and univariate (two-way ANOVA) analysis was performed considering two factors, treatment (control and stress), genotypes (pearl millet and wheat) and their interactions. Each table consists of df (degree of freedom), *F*-value (*F*-test) and *p*-value (of the *F*-test) for every factor. For K-means clustering analysis, proteins were chosen only if they were present in all four biological replicates of at least one condition. All the identified proteins were categorized into functional groups to allow a functional view of the tissue-specific proteome. The sum of the NSAF values for each functional category was then visualized using spider plots. Sparse partial

least squares (sPLS), discriminant and network analysis were performed to integrate physiological parameters and proteome data to show the interaction between proteins (predictors) and physiology (response). sPLS was performed employing R package mixOmics. Generated networks were visualized and filtered (only edges equal or higher than 0.9 were maintained) in Cytoscape v.2.8.3 (Escandon et al., 2017).

The Venn diagrams were produced using GeneVenn¹. A protein was considered as differentially expressed between two samples if three conditions were met: (1) the protein was detected in all four replicates at least in one of the treatments, (2) *p*-value for differential expression was ≤ 0.01 and (3) the fold change in protein NSAF values between the samples was at least 1.5. Volcano and spider plots were produced using Microsoft Excel 2015. Box plots were constructed using program R (version 3.5, R Core Team 2019) (package ggplot2).

RESULTS

Genotypic Variation of Physiological Responses, Plant Biomass, and Yield Under Drought Stress

In order to investigate the physiological basis of genotypic variation under drought stress, several parameters were determined, including stomatal conductance, F_v/F_m , root growth and others (see section “Materials and Methods”). All the recorded observations are provided in **Supplementary Table S2**.

Principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) was performed considering all the factors and variables of the physiological data (**Supplementary Figure S1A**). A clear separation was observed between pearl millet and wheat genotypes on discriminant function 2 (PC2) as well as between control and stress condition of each genotype on discriminant function 1 (PC1). Several physiological parameters separated individual genotypes. The data revealed that the growth parameter of all the compared genotypes was affected differently. The PLS-DA emphasizes the differential diagnostic values discriminating between the genotypes and treatments. Stomatal conductance and photosynthetic parameters such as F_v/F_m and chlorophyll measurements (SPAD) discriminated most between control and stressed plants, whereas the number of the panicles and spikelets was the highest discriminant value between the genera *Pennisetum* and *Triticum* as observed in the PCA (**Supplementary Figure S1B**).

The determination of basic growth parameters is essential in the characterization of drought stress response mechanisms (Jones, 2007). Overall, plant biomass was decreased under stress condition in all the four genotypes compared to controls (**Supplementary Table S2**). The highest reduction in biomass was observed in TRI 5630, followed by ICTP8203 (**Supplementary Table S2**), but they also showed the highest seed yield under drought. Stressed plants of the wheat genotypes have shown decrease in the seed weight compared to the pearl millet

¹<http://genevenn.sourceforge.net/>

genotypes and the most severe effect was observed in White Fife which is the most susceptible genotype to drought stress (Supplementary Table S2).

Contrasting Genotype Effects of Stomata Responses to Drought Stress

The immediate response of plants under drought stress is stomatal closure to prevent water loss via transpiration (Buckley, 2019). Plants grown under drought conditions tend to have lower stomatal conductance, thus helping to conserve water and maintain an adequate leaf water status but at the same time reducing leaf internal CO_2 concentration and photosynthesis. The precise relationship is also dependent on other factors, like genotypes, drought history and environmental conditions. We measured stomatal conductance (gs) at regular intervals from the start until the conclusion of the stress treatment, considering fully grown leaves using a PWMR-4 porometer. It was observed that control plants of tolerant genotypes TRI 5630 and ICTP8203 showed stomatal conductance ranging between 130 and $150 \text{ mmol m}^{-2}\text{s}^{-1}$ and sensitive genotypes (White Fife and 843-22B) showed conductance between 40 and $50 \text{ mmol m}^{-2}\text{s}^{-1}$ (Figure 1B and Supplementary Table S2). Stomatal conductance of the stressed plants declined as drought stress progressed. In the sensitive genotypes White Fife and 843-22B, the conductance declined gradually and reached $10 \text{ mmol m}^{-2}\text{s}^{-1}$. Interestingly, the most significant and rapid effect was observed in the tolerant genotypes TRI 5630 and ICTP8203 where stomatal conductance reached 20 and $40 \text{ mmol m}^{-2}\text{s}^{-1}$, respectively, at the end of drought treatment (Figure 1B).

Contrasting Regulation of Chlorophyll Content and F_v/F_m Under Drought Stress Revealed a Stay-Green Phenotype vs. Senescence Phenotypes in Pearl Millet and Wheat

In order to understand the photosynthetic capabilities of pearl millet and wheat genotypes, leaf chlorophyll content was determined using a SPAD meter (Figure 1C and Supplementary Table S2). The SPAD values in the stressed plants declined as the drought stress progressed. In the control condition, SPAD value ranged between 20 and 40% approximately for both the genotypes of pearl millet and wheat. Under dehydration state, SPAD value reduced to 10–24% approximately. The highest chlorophyll content recorded during stress was in ICTP8203 (24%), followed by 843-22B (14%), TRI 5630 (14%), and White Fife (>10%).

Drought stress consistently and significantly reduced the maximum efficiency of PSII photochemistry (F_v/F_m), though this effect varied in its severity among the different genotypes (Figure 1D and Supplementary Table S2). By the end of the drought regime, the highest F_v/F_m ratio was observed in ICTP8203, followed by 843-22B and TRI 5630. This correlates directly with the SPAD measurements, indicating the stay-green phenotype of ICTP8203 in contrast to the other genotypes. The lowest reading was recorded in White Fife (Figure 1D).

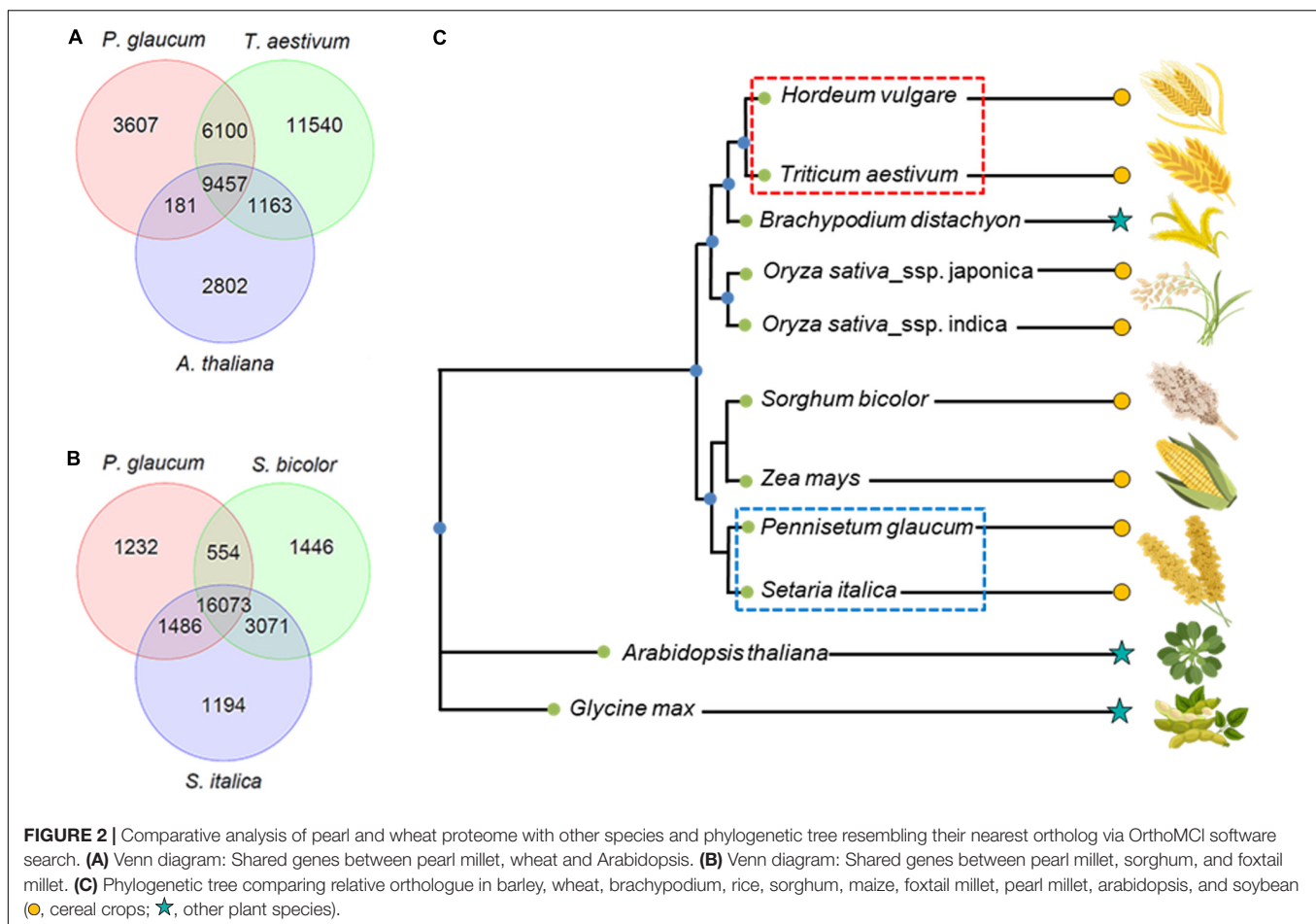
Comparative Analysis of Tissue-Specific Pearl Millet and Wheat Drought Stress Proteomes

To perform a detailed proteome study, not only a full genome sequence is required, but also accurate gene annotation plays a critical role (Valledor et al., 2012). Here, for a comparative proteomics study, we identified unique and shared gene families between pearl millet and wheat using OrthoMCL (Li et al., 2003). Because of the high redundancy of the wheat genome annotation, the wheat sequences were clustered using CD-HIT (Fu et al., 2012), resulting in a set of 69,215 sequences (from the initial 136,866). From the total 46,954 gene families detected by the OrthoMCL algorithm, 9,457 were found to be shared between pearl millet, wheat and *Arabidopsis* (Figure 2A). On the gene level, ~40% of the proteomes was shared between *Arabidopsis* and pearl millet (13,710 proteins for *Arabidopsis* and 11,535 proteins for pearl millet). Comparative analysis was also performed considering sorghum (*Sorghum bicolor*) and foxtail millet (*Setaria italica*) (Figure 2B). In the comparison between pearl millet and sorghum, 70–75% of the proteome was shared (19,865 proteins for pearl millet and 20,279 proteins for sorghum). Foxtail millet represents the evolutionarily closest plant species to pearl millet in this ortholog family analysis (Figure 2C, blue). Between these two species, ~70–80% of the proteome was shared (21,447 proteins for pearl millet and 22,170 proteins for foxtail millet). For wheat, the closest related plant species is barley (*Hordeum vulgare*) (Figure 2C, red), where coverage of the former is 59% (30,624 proteins of the reduced redundancy sequence set) and 94% for the latter (18,239 proteins).

From all the detected peptides in roots, leaves and seeds, 12,558 proteins were identified from both pearl millet and wheat genotypes, of which 4,564 proteins were identified in pearl millet (843-22B and ICTP8203) genotypes (Supplementary Table S3). In wheat genotypes (White Fife and TRI 5630), 7,994 proteins were identified (Supplementary Table S4). In order to generate a broad survey of identified proteins with altered tissue-specific abundance under drought stress, a Venn analysis was conducted which determines the dynamics of the proteome in selected genotypes of pearl millet and wheat under control and drought stress (Supplementary Figure S2). We performed two-way ANOVA analysis of the identified pearl millet and wheat proteome for every factor: (i) treatment (control and stress) and, (ii) genotypes (pearl millet and wheat) and (iii) their interactions. Here, each tissue (root, leaf and seed) was analyzed separately (Supplementary Table S5).

Functional Categorization and Statistical Analysis of the Drought Stress Proteome in Contrasting Pearl Millet and Wheat Genotypes

The tissue-specific DEPs (differentially expressed proteins) in pearl millet (843-22B and ICTP8203) and wheat (White Fife and TRI 5630) genotypes are represented using volcano plot's (Figures 3A,B), and the list of DEPs are summarized in

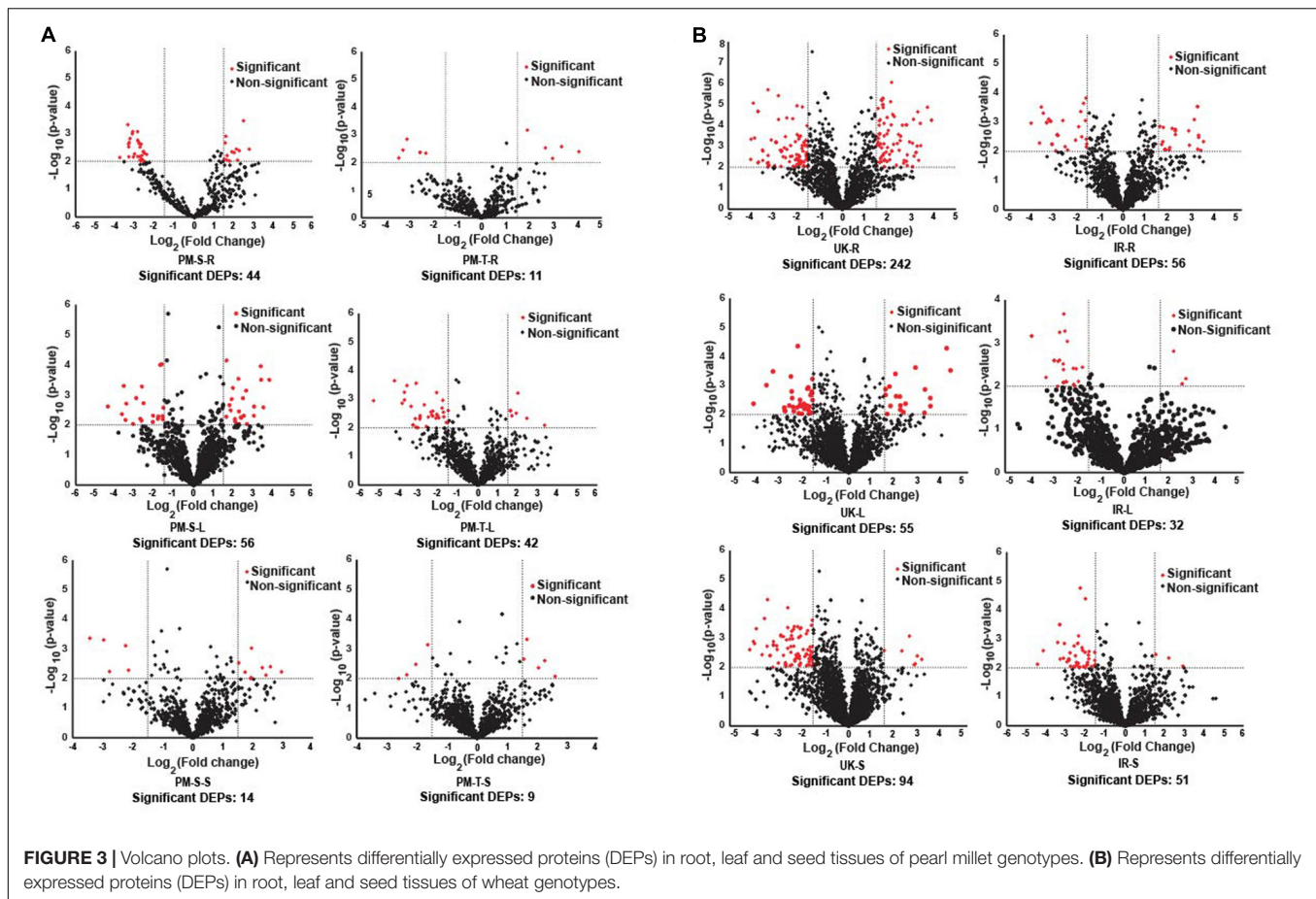


Supplementary Tables S6, S7. K-means clustering analysis was employed to investigate the co-expression/abundance pattern of the identified proteins from the compared genotypes. For cluster analysis, proteins were considered if they were present in all the four biological replicates of at least in one tissue/condition. Tissue-specific grouping of proteins in different condition (control and stress) lead us with clusters $k = 30$ in 843-22B and $k = 35$ in ICTP8203 (**Supplementary Table S8**), similarly, $k = 50$ in White Fife and TRI 5630 (**Supplementary Table S9**). Cluster analysis revealed specific groups of proteins with changing abundance in tissue or drought stress condition.

Functional categorization of the identified proteins was performed according to Ramsak et al. (2014). Tissue-specific functional distribution of the proteome for pearl millet (843-22B and ICTP8203) and wheat (White Fife and TRI 5630) genotypes under drought stress is depicted in **Figure 4** via spider plots using the total NSAF score summed up for different functional categories (Chaturvedi et al., 2013, 2015). The overall pattern demonstrates a remarkable variation of proteome functionality between the sensitive and tolerant-genotypes expressed as ratios of drought stress vs. controls. In pearl millet, major enhanced functional categories in 843-22B are transport and stress-related proteins in the root, mitochondrial electron transport, TCA cycle, C1-metabolism in leaf. In contrast, ICTP8203 showed drought

stress enhanced protein functions for cell wall degradation, signaling and polyamine metabolism in the root, light reactions, photorespiration, transport and signaling in leaf and a strong regulation was observed in the development and polyamine metabolism in seed tissue. In wheat, White Fife showed enhanced regulation in protein categories of cell wall synthesis, mitochondrial electron transport, and redox in root tissue. In the leaf, TRI 5630 showed pronounced proteome regulation in the functional categories of gluconeogenesis, lipid metabolism, amino acid metabolism and carbohydrate metabolism compared to the White Fife. Similarly, seed proteome of TRI5630 showed enhanced proteome regulation compared to White Fife, e.g., C1 metabolism, secondary metabolism and transport (**Figure 4**). Functional categories and related proteins distinguishing the genotype- and tissue-specific drought stress response according to **Figure 4** can be found in **Supplementary Tables S10, S11**.

Principal component analysis (PCA) of protein NSAF scores was performed using COVAIN (Sun and Weckwerth, 2012). All the tissues were separated by the first principal component (PC1) providing hints for tissue-specific proteome functionality in both pearl millet and wheat. In pearl millet, the strongest variation was observed in root and leaf tissues (**Supplementary Figure S3A**). In PCA of ICTP8203, leaf proteome showed the strongest variation in response to drought stress compared



to roots and seeds (**Supplementary Figures S3B,C**). Positive loadings of PC1 represent proteins with higher abundances in roots, whereas negative loadings depicted higher levels in leaf and seed tissues (**Supplementary Table S12**). In 843-22B, the root proteome showed the most substantial variation followed by seed and leaf tissues in response to drought stress (**Supplementary Figure S3D**). These tissue-specific proteome effects were in contrast to the tolerant genotype (ICTP8203). Considering loadings, the highest positive loading showed proteins with higher abundance in seed and leaf tissues, while negative loading showed proteins with higher abundance in root tissue (**Supplementary Figure S3E**). Interestingly, it was observed that the seed proteome of both pearl millet genotypes showed only a small difference between control and stress condition.

Contrasting proteome effects were observed in wheat genotypes compared to pearl millet genotypes. Here, the strongest variation was observed in seed and leaf tissues (**Supplementary Figure S4A**). In TRI 5630, the seed proteome showed the strongest variation in response to drought stress compared to root and leaf tissues (**Supplementary Figure S4B**). Positive loadings of PC1 represent proteins with higher abundances in seed and leaf tissues, whereas negative loadings depicted higher levels in root tissue (**Supplementary Figure S4C** and **Supplementary Table S13**). However, in wheat sensitive

genotype White Fife, most substantial variation was observed in the leaf followed by root tissues (**Supplementary Figures S4D,E**).

DISCUSSION

Physiological Comparison of Whole Plant Responses to Drought Stress

To integrate all physiological information into an intuitive coherent visualization model, we used the visualization strategy of Odum, an approach which integrates systems-theoretical ideas for the analysis of multivariate systems in ecology (Odum, 1994; Weckwerth, 2019). Here, each symbol and size determine systems state variables and their quantity representative for the individual pearl millet and wheat phenotypes under drought stress (**Figure 5**). Using this visualization strategy an n-dimensional multivariate data matrix and its intrinsic dynamics can be intuitively recognized by visual inspection. This principle is also known as coherent perception, e.g., face recognition, and by using symbols and sizes, we translate highly complex multivariate data into an intuitive visual model otherwise depicted by multivariate statistics such as PCA (Weckwerth and Morgenthal, 2005; Weckwerth, 2008; Weckwerth, 2019). **Figures 5A,B** determines the physiological response of the wheat and pearl millet under well-watered condition. Grain yield allows direct estimation

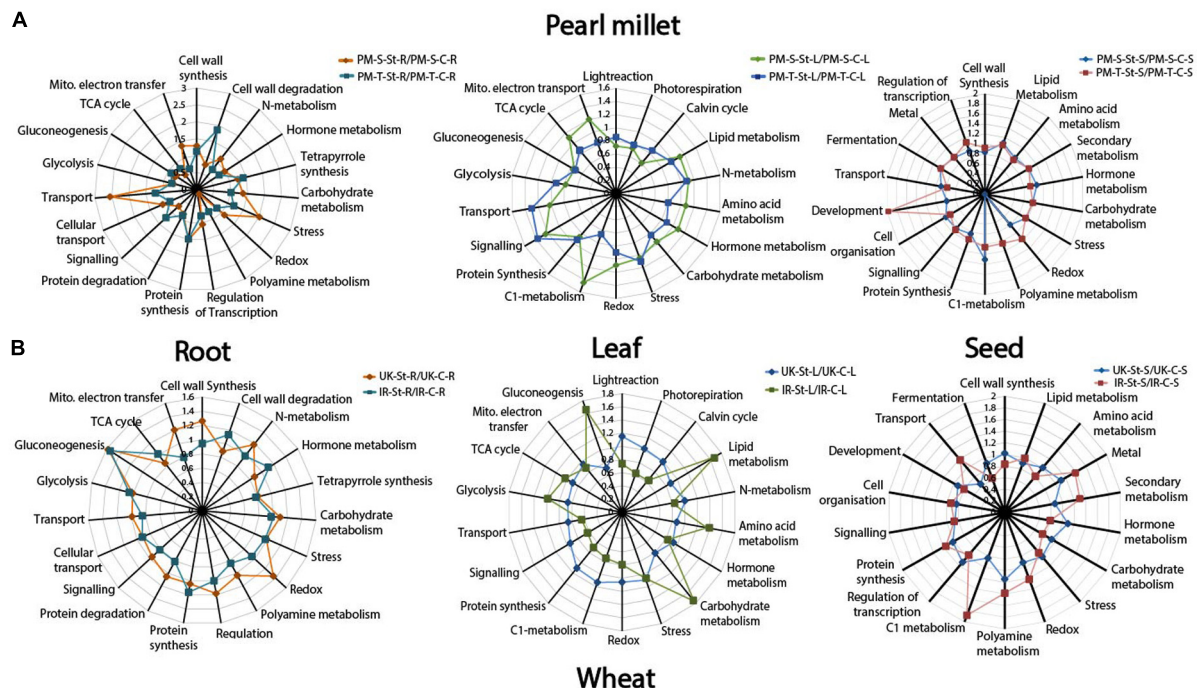


FIGURE 4 | Spider plots. **(A)** Represents functional distribution of the proteome in the root, leaf, and seed tissues of pearl millet genotypes. **(B)** Represents functional distribution of the proteome in the root, leaf, and seed tissues of wheat genotypes.

of the drought tolerance capacity of the individual genotypes (Fischer and Wood, 1979). Under well-watered conditions, pearl millet (843-22B and ICTP8203) and wheat (White Fife and TRI5630) genotypes were found to have comparable grain yields. By contrast, water stress treatment resulted in different grain yields between the genotypes. A significant response was observed between White Fife and TRI5630. White Fife was not able to maintain its yield under drought stress (Figures 5C,E and Supplementary Table S2). A similar response was also observed by Inzanloo and co-workers, where a sensitive genotype Kukri showed a significant drop in the grain yield under stress condition (Inzanloo et al., 2008).

In contrast, pearl millet genotypes were able to restore their grain yield under drought stress (Figures 5D,F). Bidinger et al. (1987) reported a similar response of pearl millet under mid-season drought (panicle initiation to flowering) stress. This effect is also related to biomass production during drought stress. The significant reduction in biomass of these drought-tolerant genotypes and the resulting higher harvest index can be attributed as an adaptive response where plants endure low tissue water content through maintenance of cell turgor via osmotic adjustment and cellular elasticity and divert their entire energy to protect seed production under harsh conditions (Farooq et al., 2009). Also at the proteome level, there are relatively small changes in both pearl millet genotypes between well-watered and drought stress conditions which are in stark contrast to wheat. Accordingly, the proteome data reflect the physiological data. In this context, it is being reported by Begg that pearl millet even under favorable conditions tends to have a shorter

crop cycle than any other cereals because it has a “built-in” drought escape mechanism of early flowering, inherited from its wild progenitors which are evolved in semi-desert environments. Therefore, pearl millet has not only a short crop cycle but also short grain-filling period and small seed sizes which is a clear advantage in unfavorable growth conditions such as heat and drought (Begg, 1965).

One significant difference in C_3 and C_4 plants species is their photosynthetic capacities. Globally, 85% of higher plant species follow C_3 -type photosynthesis whereas only 4% of the plant species belong to the C_4 -type majorly originated in arid regions where high temperature occurs with water stress (Yamori et al., 2014). To investigate these photosynthetic capabilities in more detail chlorophyll content and chlorophyll fluorescence of plants were measured along with stomatal behavior (see below). In all the genotypes, the imposition of drought stress resulted in a decrease in chlorophyll content (Figure 1C and Supplementary Table S2). In this study, the pale leaves with a lower chlorophyll content in White Fife (> 10% SPAD units), 843-22B (14% SPAD units) and TRI 5630 (14% SPAD units) senesced early, while the green leaves with a high chlorophyll content in ICTP8203 consistently stayed green. The similar effect was observed in drought-tolerant and susceptible cultivars of peanut under water stress (Katam et al., 2016). Previous studies also revealed that a decrease in chlorophyll concentration under drought stress could be related to the increase in the activity of enzyme chlorophyllase (Ashraf et al., 1994). Drought stress-induced decrease in the pigment content was also previously reported in several plant species, including durum wheat (Loggini et al., 1999). The

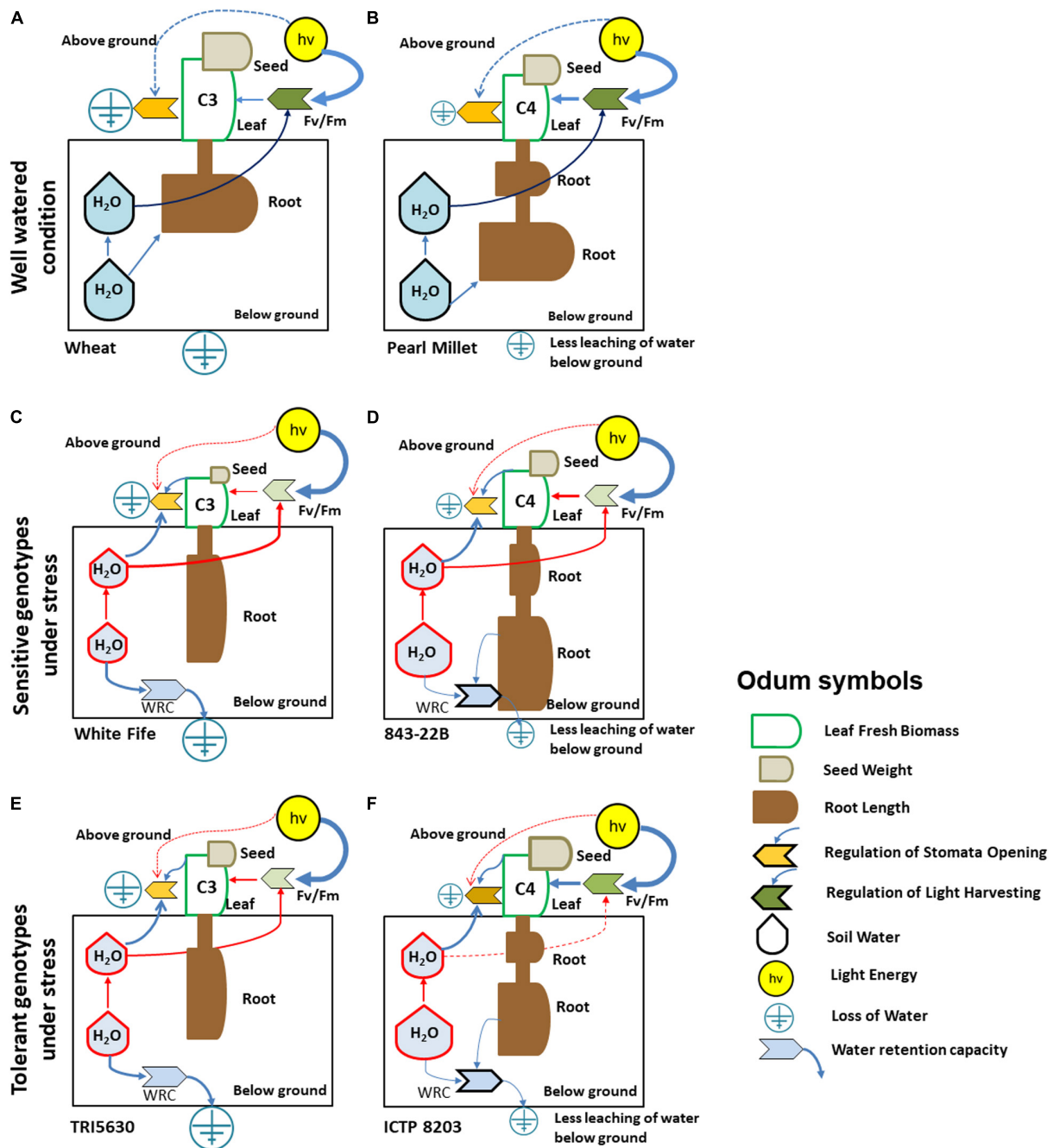


FIGURE 5 | Model for significant and relevant eco-physiological responses of pearl millet and wheat genotypes under drought stress using Odum's symbols. (A) Model for C₃ wheat under well-watered condition. (B) Model for C₄ pearl millet under well-watered condition. (C) Model for White Fife genotype under stress condition. (D) Model for 843-22B genotype under stress condition. (E) Model for TRI 5630 genotype under stress condition. (F) Model for ICTP8203 genotype under stress condition. [Different color code indicates different regulation of the physiological parameters under control (dark color) and stress (light color, diminished regulation; semi-dark color, moderate regulation) condition; Blue line indicates physiological mass/energy transfer; Red line indicates altered physiological response under drought stress; Root length: (A,B) feature short, dense roots, (C,E) feature moderate/long length and not so dense roots, (D,F) exhibit moderate/long and dense roots].

stay-green trait, in contrast, protects the leaf from the degradation of chlorophyll, stabilizes photosystem and helps to produce normal grain (Thomas and Ougham, 2014; Kamal et al., 2019). However, few reports are available on the mechanism of how this

trait protects chlorophyll under drought and which dominant genes control this trait under drought conditions (Walulu et al., 1994). Senescence is typically characterized by chlorophyll loss and a progressive decline in photosynthetic capacity. Early onset

of senescence affects assimilation and grain filling in crop plants (Xu et al., 2000). This effect can be correlated to wheat genotypes as they were not able to maintain their chlorophyll content under drought. Accordingly, ICTP8203 showed a stay-green phenotype in contrast to all the other genotypes (Thomas and Howarth, 2000). In this context, F_v/F_m ratio was also highest in ICTP8203 (Figure 5E) followed by 843-22B and TRI 5630 under drought treatment (Figure 1D and Supplementary Table S2). The F_v/F_m ratio of White Fife (Figure 5C) was significantly reduced, indicating a severely impaired photosystem under drought conditions. Programmed leaf senescence is initiated contributing to the plant survival under drought conditions (Lu and Zhang, 1998; Munne-Bosch and Alegre, 2000, 2004; Lu et al., 2002) but also resulting in yield losses (Borrell and Hammer, 2000; Jiang et al., 2004; Rivero et al., 2007).

The tolerant genotypes of pearl millet and wheat ICTP8203 and TRI 5630 also showed a different stomata regulation than the susceptible ones. ICTP8203 and TRI 5630 showed stomatal conductance ranging between 130 and 150 $\text{mmol m}^{-2}\text{s}^{-1}$ under control condition, already higher than in the sensitive genotypes (White Fife and 843-22B) ranging from 40 to 50 $\text{mmol m}^{-2}\text{s}^{-1}$. During drought stress, the range of stomata closure was higher and more rapid in the tolerant genotypes. A similar response was observed in the study performed on soybean genotypes (Liu et al., 2005) and *Amaranthus* species under drought stress (Liu and Stutzel, 2002). The rapid stomatal response may act as a drought resistance mechanism, which permits to keep water for later use and thus maintain higher leaf water potentials (Jones, 1974). In principle, stomatal closure protects plants against excessive water loss but also restricts the diffusion of CO_2 into the photosynthetic parenchyma, especially for C_3 plants. Stomatal closure causes more significant decrease in transpiration than in photosynthesis rates, thereby increasing the relative leaf water use efficiency (WUE) (Pou et al., 2008). Thus, more dynamic and more extensive regulation of stomata in the tolerant genotypes is one of the pre-requisites for better performance under drought stress. Furthermore, the stomatal limitation on photosynthesis can be accompanied by a decrease in the rate of consumption of ATP and NADPH for CO_2 assimilation that could result in a decrease in the rate of linear electron transport and consequently in F_v/F_m (Baker and Rosenqvist, 2004) which was primarily observed in the susceptible wheat genotype, White Fife.

Considering the observations of photosynthetic capabilities and stomata regulation the tolerant varieties seem to compensate differences between C_3 - and C_4 -type photosynthesis. The differences are instead found in the stay-green vs. programmed leaf senescence phenotypes.

Another very strong effect is the different root length between wheat and pearl millet but also between the intraspecific genotypes (Figure 5 and Supplementary Figure S5). Due to this difference in root length, water retention capacity is very different for wheat and pearl millet genotypes (Figures 5C–F). Root length was increased in all the genotypes under drought stress (Figures 5C–F). Root length appeared to be an important trait for drought stress tolerance, as reported in the previous study (Leishman and Westoby, 1994). However, unexpectedly, here the sensitive pearl millet genotype 843-22B showed maximum

root length as compared to other genotypes. The impact of root system and its mechanism on yield under drought conditions is also comprehensively discussed in many major crops (Tuberosa and Salvi, 2006; de Dorlodot et al., 2007; Comas et al., 2013; Sehgal et al., 2015). Controlled greenhouse and field conditions show different variations of plant functional and molecular traits (Hoeftwarter et al., 2008; Holmgren et al., 2012; Nagler et al., 2018; Weckwerth et al., 2020). In future studies all the drought-related traits and molecular signatures which are described in our study will be also tested under field conditions.

Proteome Signature for “Stay-Green” and “Senescence” Trait Under Drought Stress

Stay-green is an important agronomical trait which can contribute to higher yield production under drought stress condition (Harris et al., 2007; Thomas and Ougham, 2014). However, not much is known about the protein changes leading to this effect. In the present study, a significant change in protein patterns of pearl millet genotypes provided a clear indication of the processes that underlie the stay-green or senescence trait in ICTP8203 and 843-22B, respectively (Figure 6A and Supplementary Table S14). ICTP8203 showed enhanced regulation in photosynthetic activity under drought stress. This correlates with significant higher levels of chlorophyll a-b binding protein (Pgl_GLEAN_10021964), protein kinase (Pgl_GLEAN_10013653), thylakoid lumenal 19 kDa protein (Pgl_GLEAN_10006356), ferredoxin-NADP reductase (Pgl_GLEAN_10033849). This higher photosystem activity resulted in lower levels of reactive oxygen species (ROS) proteins such as peroxidases (Pgl_GLEAN_10014871, Pgl_GLEAN_10027105, Pgl_GLEAN_10006633), glutathione reductase (Pgl_GLEAN_10019381, Pgl_GLEAN_10036180), glutathione synthetase (Pgl_GLEAN_10035689), and peroxiredoxin (Pgl_GLEAN_10024324) under drought stress. Furthermore, the higher levels of several stay-green-associated proteins such as 14-3-3 (Pgl_GLEAN_10007318), chlorophyll synthesis proteins (such as magnesium chelatase ATPase subunit I) (Pgl_GLEAN_10038264), ribulose biphosphate carboxylase small chain (Pgl_GLEAN_10020566), and uroporphyrinogen decarboxylase (Pgl_GLEAN_10000112) were also observed in ICTP8203. 14-3-3 are the binding proteins that show strong interaction with the enzymes involved in nitrogen and carbon metabolisms which may influence the degradation process (Huber et al., 1996). Overexpression of *Arabidopsis* gene GF14< (which encodes 14-3-3 protein) in cotton lead to stay-green phenotype and also improved drought tolerance of transgenic cotton under drought stress (Yan et al., 2004). Similarly, overexpression of 14-3-3 protein delayed leaf senescence in potato plant (Wilczynski et al., 1998). A protein related to photorespiration (aminomethyltransferase; Pgl_GLEAN_10027187) also showed higher levels in ICTP8203 under drought stress (Supplementary Table S14).

Interestingly, we observed decreased levels of ferrochelatase (Pgl_GLEAN_10011603) which reduces cytotoxicity and in

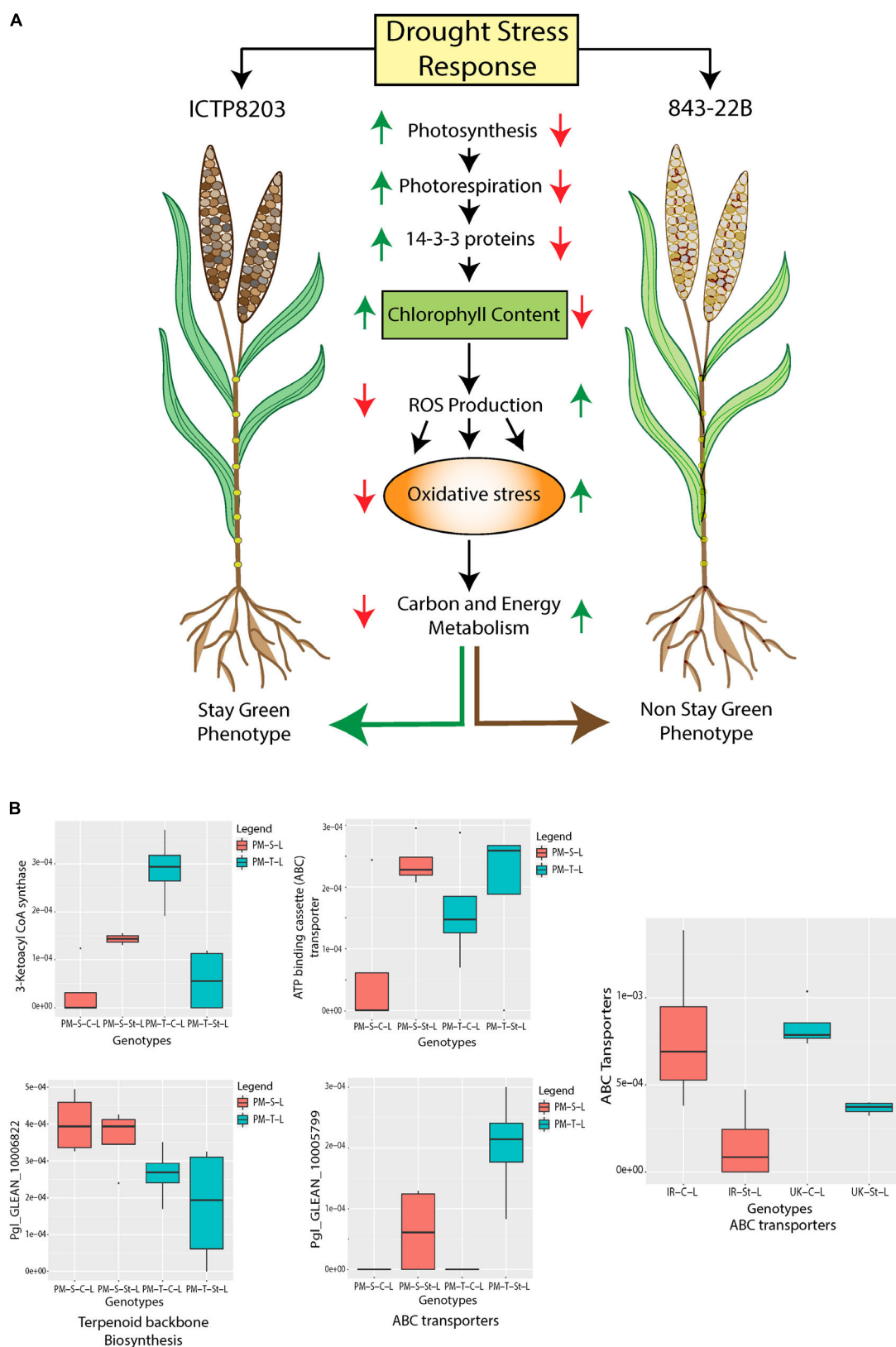


FIGURE 6 | (A) Regulation of proteome underlying “Stay-green” trait in pearl millet genotypes (843-22B and ICTP8203). **(B)** Regulation of wax biosynthesis proteins in pearl millet and wheat genotypes under control and stress condition.

turn, increases chlorophyll biosynthesis (Nagahatenna et al., 2015). In contrast, the sensitive pearl millet genotype 843-22B demonstrated opposite regulation of the proteome compared to ICTP8203, contradicting the stay-green process (**Figure 6A**). Proteins binned into the functional category of RING finger ubiquitin showed increased levels in 843-22B compared to ICTP8203 under drought stress. The observed regulation of the proteome is positively correlated to the physiological analysis where maximum efficiency of PSII photochemistry (F_v/F_m), chlorophyll content and yield was highest in ICTP8203 by the end of the drought stress. Hence, the proteome of ICTP8203 can be identified as “stay-green” signature.

Wheat genotypes demonstrated a different regulation of the proteome compared to pearl millet and did not show the “stay-green” trait at the phenotypic level. The significant changing pattern of proteomes, indicated “senescence” processes (non-sequential and sequential) in TRI5630 and White Fife, respectively (**Supplementary Table S14**). TRI5630 showed a significant increase in sucrose synthase 4 (SS) (W5I774, A0A1D6SCX5) (more than 6 and 11-fold change, respectively) under drought stress. Increased levels of sucrose content in leaf contributed to non-sequential senescence process. Moreover, increased sugar mobilization promotes stable yield and enhances the tolerance mechanism, which is indicated by higher stem reserve remobilization under drought stress (Shi et al., 2016). Remobilization of stored carbon reserves in wheat is also facilitated under drought stress, which enhances plant senescence process and accelerates grain filling process (Yang et al., 2000, 2001). This can also be correlated to the phenotypic data of TRI5630 showing stable yield despite decreased photosynthetic and chlorophyll biosynthetic proteins.

Interestingly, we also observed increased levels of 4- α -glucanotransferase (W5BL76, W5B4C2) (with > 2-fold change) in TRI5630 under drought stress whereas in White Fife these isoforms showed decreased levels or no change under stress. It has been reported that 4- α -glucanotransferase is an essential component of the pathway from starch to sucrose and cellular metabolism in leaves at night (Chia et al., 2004). Therefore, increased levels of 4- α -glucanotransferase can indicate higher sucrose content in the leaf of TRI5630, which supports the leaf non-sequential senescence process. In White Fife, despite higher levels of photosynthetic and chlorophyll biosynthetic proteins and a decrease in reactive oxygen (ROS) proteins, the total yield was decreased under stress condition (**Supplementary Table S14**). The possible reason could be a sequential senescence process in which young leaves are successively formed at the top region of the plant and lower older leaves undergo senescence.

Regulation of Wax Biosynthetic Proteins Under Drought Stress Indicate Higher Capacities for Drought Protective Cuticular Wax Biosynthesis in Pearl Millet Compared to Wheat

Plants are evolved with diverse adaptive strategies to cope with water deficit conditions. Accumulation of cuticular waxes

is such a strategy which contributes to drought resistance (Samuels et al., 2008; Lee and Suh, 2015). Seo et al. (2009) demonstrated that increase in the cuticular wax synthesis improves drought tolerance in Arabidopsis species. Reduced wax production leads to drought sensitivity in rice (Zhu and Xiong, 2013). It is also known that cuticular wax biosynthesis is also controlled at post-transcriptional and post-translational levels (Lee and Suh, 2015). At the proteome level, we identified the regulation of two key rate-limiting enzymes of cuticular wax biosynthesis in pearl millet and wheat leaf tissue which includes 3-ketoacyl-CoA synthase (KCS) and ATP-binding cassette (ABC) transporter. In pearl millet genotypes, one of the 3-ketoacyl-CoA synthase (Pgl_GLEAN_10030730) showed enhanced levels under stress condition in the sensitive genotype 843-22B compared to the tolerant genotype ICTP8203 (**Figure 6B**). Similarly, GWAS study lead to the identification 3 SNPs located between two predicted genes encoding for 3-ketoacyl-CoA synthase in pearl millet under drought stress (Debieu et al., 2018). Interestingly in wheat genotypes (White Fife and TRI 5630), this gene family was not detected in either condition. 3-ketoacyl-CoA synthase is not only involved in decarbonylation and acyl-reduction of wax synthesis pathways but also involved in elongation of C24 fatty acids which is an essential condensation step during wax and suberin biosynthesis. A study performed by Yu and coworkers demonstrated that OsWSL1 encodes 3-ketoacyl-CoA synthase (KCS) genes in rice, catalyzes the formation of C20–C24 VLCFA precursors of leaf waxes. The OsWSL1 mutant showed a pleiotropic phenotype with decreased growth, sparse wax crystals and drought sensitivity, suggesting that OsWSL1 may be relevant to drought tolerance (Yu et al., 2008). Export of cuticular wax is mediated by the ATP binding cassette (ABC) transporters (Pighin et al., 2004; Bird et al., 2007; Panikashvili et al., 2007). In both the pearl millet genotypes three ATP-binding cassette (ABC) transporter proteins (Pgl_GLEAN_10004859, Pgl_GLEAN_10002141, Pgl_GLEAN_10006800) were identified and showed increased levels under stress condition compared to controls. Intriguingly, this protein showed an opposite regulation pattern in wheat genotypes (A0A1D5VIG5, A0A1D6C5F5, A0A1D6BMJ3, A0A1D6D783) under drought stress (**Figure 6B**). Information on wax biosynthetic genes is sparse in wheat due to the lack of functional genomic studies. However, it is known that wheat employs another parallel wax biosynthetic pathway, which is predominant in the reproductive stages and responsible for the biosynthesis of β -diketones (Tulloch, 1973). Recently, the pearl millet genome study has shown substantial enrichment and expansion of wax biosynthetic genes which may contribute to heat and drought tolerance of this crop in semi-arid regions (Varshney et al., 2017) compared to other cereals. Taking a look at the translational level of these genes in the leaf tissue of pearl millet genotypes, four genes Pgl_GLEAN_10006822, Pgl_GLEAN_10030730, Pgl_GLEAN_10005799, and Pgl_GLEAN_10005798 were identified which belonged to the group of terpenoid backbone biosynthesis, suberin biosynthesis and ABC transporters, respectively. These data indicate that wax biosynthesis is enhanced in pearl millet at the proteome level, especially in

ICTP8203. In contrast, these pathways are not detected in wheat and drought-dependent enhancement is also not observed.

Drought Responsive Regulation of the Key Photosynthetic Proteins of Pearl Millet (NAD-ME Type) and Wheat Under Drought Stress

Global depletion of atmospheric CO₂ levels led to the evolution of C₄ photosynthesis from ancestral C₃ photosynthesis. Among C₄ plants, there are three biochemical subtypes, based on the C₄ acid decarboxylation enzyme in the bundle sheath (Hatch, 1987; Leegood, 2002) (1) NADP-malic enzyme (NADP-ME) type, (2) NAD-malic enzyme (NAD-ME) type, and (3) phosphoenolpyruvate carboxykinase (PEPCK) type (Leegood, 2002). Distribution of C₄ grasses is strongly influenced by rainfall level, e.g., areas with decreasing rainfall (from 900 to 50 mm per annum) demonstrates an increased abundance of NAD-ME subtype grasses compared to NADP-ME subtype grasses. This geographical distribution of C₄ grasses with different biochemical subtypes may also reflect different drought tolerance capacities (Ghannoum, 2009). However, at the proteome level, there is no evidence suggesting that these three C₄ biochemical pathways have different sensitivities to water stress. Pearl millet has been classified as a NAD-ME subtype (Edwards and Walker, 1983). The enzymes involved in C₄ photosynthesis are also present in C₃ plants of course without the Kranz anatomy but expression regulation is different, and activities are much lower. These enzymes operate for different metabolic processes, and they also have different inter- and intracellular localization. At the proteome level, we have identified all significant enzymes associated with the C₄ pearl millet NAD-ME subtype

photosynthesis under control and drought stress in both the genotypes wheat with different abundance level (**Figure 7** and **Supplementary Figure S6**).

Carbonic anhydrase (CA) and phosphoenolpyruvate carboxylase (PEPC) are two important enzymes at the beginning of the C_4 carbon fixation process and may be directly related to photosynthesis efficiency. They both are located in the cytosol of mesophyll cells, whereas in C_3 plants, they are predominantly located in chloroplast stroma (Ignatova et al., 2019). CA at the proteome level showed decreased levels under stress condition compared to controls in both the genotypes of pearl millet (Pgl_GLEAN_10019649, Pgl_GLEAN_10007313, Pgl_GLEAN_10019645). In contrast, PEPC (Pgl_GLEAN_10033512, Pgl_GLEAN_10037989, Pgl_GLEAN_10033055, Pgl_GLEAN_10026714, Pgl_GLEAN_10036281) showed increased levels under drought stress in the tolerant genotype ICTP8203 (**Figure 7**). In wheat genotypes, CA (A0A1D5WHU7) showed increased levels in White Fife under stress condition but did not show any regulation in TRI 5630. PEPC (A0A1D5YFR1, A0A1D6RUC3, A0A1D6C446, A0A1D6BRX9, A0A1D6RR40, A0A1D6ALV9, A0A1D5WD42, and A0A1D6AHU2) showed decreased levels in both the genotype under stress condition (**Figure 7**). Phosphoenolpyruvate carboxylase (PEPC), the second most abundant enzyme of the C_4 pathway, is an important and multifaceted enzyme that catalyzes the irreversible β -carboxylation of phosphoenolpyruvate (PEP) to yield oxaloacetate (OAA) and inorganic phosphate (Pi) (Hatch, 1987). It also provides oxaloacetic acid to the tricarboxylic acid cycle (TCA cycle) and catalyzes reactions involved in amino acid metabolism in C_3 plants and non-photosynthetic tissues of C_4 plants (Westhoff and Gowik, 2004). Ding and co-workers

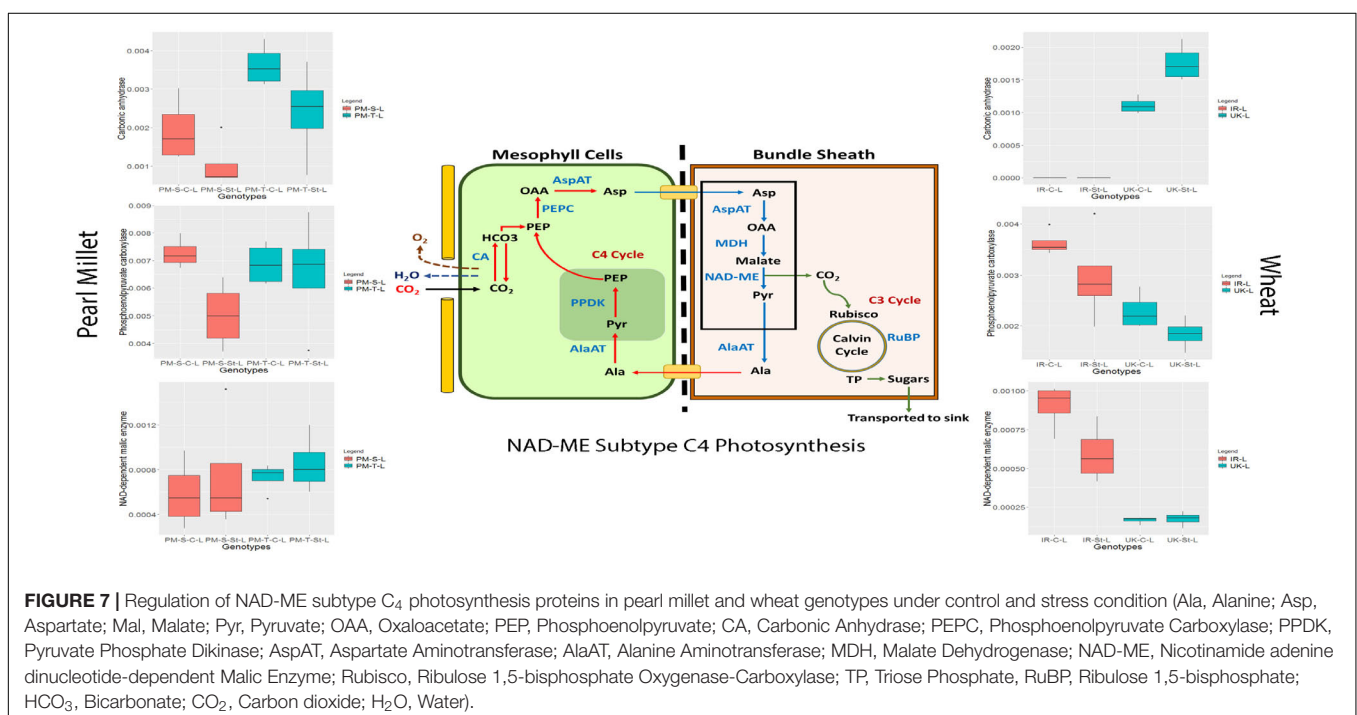


FIGURE 7 | Regulation of NAD-ME subtype C₄ photosynthesis proteins in pearl millet and wheat genotypes under control and stress condition (Ala, Alanine; Asp, Aspartate; Mal, Malate; Pyr, Pyruvate; OAA, Oxaloacetate; PEP, Phosphoenolpyruvate; CA, Carbonic Anhydrase; PEPC, Phosphoenolpyruvate Carboxylase; PPDK, Pyruvate Phosphate Dikinase; AspAT, Aspartate Aminotransferase; AlaAT, Alanine Aminotransferase; MDH, Malate Dehydrogenase; NAD-ME, Nicotinamide adenine dinucleotide-dependent Malic Enzyme; Rubisco, Ribulose 1,5-bisphosphate Oxygenase-Carboxylase; TP, Triose Phosphate, RuBP, Ribulose 1,5-bisphosphate; HCO₃⁻, Bicarbonate; CO₂, Carbon dioxide; H₂O, Water).

demonstrated that overexpression of PEPC led to higher CO₂ assimilation compared to wild-type under progressive drought stress (Ding et al., 2015). Similarly, Jiao et al. (2002) determined that expression of maize C₄-PEPC in transgenic rice led to improved photodetoxication and photosynthetic capacity under drought stress. However, mutation of two PEPC genes of Arabidopsis [PPC1 (AT1G53310), PPC2 (AT2G42600)] led to severe growth arrest phenotype and reduced the synthesis of malate and citrate and severely suppressed ammonium assimilation (Shi et al., 2015).

Interestingly, NAD-malic enzyme (NAD-ME) showed increased levels in stressed plants compared to controls in both the genotypes of pearl millet (Pgl_GLEAN_10013928, Pgl_GLEAN_10034558). In contrast, the reverse trend was observed in wheat genotypes (A0A1D5S6L5, A0A1D5U9X7) under drought stress (Figure 7). NAD-ME plays a major role in determining flux through the TCA cycle by providing pyruvate for oxidation. However, studies claim that antisense potato lines do now show any perturbation in flux through the TCA cycle but an alteration in glycolytic metabolism (Jenner et al., 2001).

Several other enzymes were also identified which included malate dehydrogenase (MDH), pyruvate phosphate dikinase (PPDK), aspartate aminotransferases which are majorly located in mitochondria of bundle sheath and cytosol of mesophyll cells in C₄ plants (Taniguchi et al., 1995), and alanine aminotransferases that lead to the reversible conversion of pyruvate to alanine. All the identified proteins showed reduced levels under stress condition compared to controls in pearl millet genotypes with a significant response in 843-22B (Supplementary Figure S6). In contrast, in wheat genotypes, only the sensitive genotype White Fife showed decreased levels of these proteins compared to controls. In contrast, the tolerant genotype TRI 5630 showed enhanced regulation under stress condition. These identifications indicate that the combined activities of several proteins may enable tolerant genotypes ICTP8203 and TRI 5630 to retain its photosynthetic efficiency under drought stress compared to sensitive genotypes 843-22B and White Fife. However, the pattern does not lead to a clear functional interpretation, and future studies will focus more on these pathways.

Integration of Proteome With Physiological Data of Pearl Millet and Wheat Phenotypes Demonstrates the Enormous Plasticity of Drought Stress Adaptation

The physiological data of the preceding discussions were used as predictable traits to identify protein networks which show high predictive power. Accordingly, we had several scenarios where we could exploit the predictive power of protein correlation networks for traits such as yield (seed weight), harvest index, root length and many more. For statistical modeling, we employed a method called sparse least square discriminant analysis sPLS-DA (see section “Materials and Methods”). Plants use different strategies to cope with drought stress, ranging from drought avoidance to desiccation. Drought avoidance is associated with

the minimization of water loss and simultaneously maximization of water uptake (Ludlow and Muchow, 1990). Evidence suggests that the adverse effects of drought can be successfully avoided by changing carbon allocation patterns to allow the formation of a deep root system before the onset of a growth-limiting water shortage. Accordingly, root traits are now considered as important targets under drought stress for yield improvement. Therefore, we decided to identify protein correlation networks predictive for root length which has been highly distinguishable between pearl millet and wheat genotypes. Pearl millet genotypes showed increased root length under stress condition compared to controls with extensive elongation in 843-22B. Using sPLS, key stress proteins correlating with the root length were identified for each genotype individually (Supplementary Figure S7 and Supplementary Table S15). The identified proteins were functionally classified using MAPMAN and represented in an interacting network based on the color of the functional bins (Supplementary Table S15).

Furthermore, the discriminant analysis allowed determining differences of predictive protein levels between 843-22B and ICTP8203 for root length (Figure 8A). The genotypes showed remarkable plasticity and differences in the same set of predictive protein correlation networks (Figure 8A). Based on this analysis, proteins binned in the functional categories of stress, reactive oxygen species (ROS) and oxidative pentose phosphate pathway (OPP) showed the highest correlation scores with the root length. Annexin protein was positively correlated to the 843-22B (Figure 8A). Plant annexins are Ca²⁺ dependent phospholipid-binding proteins, and they participate in the regulation of plant development as well as in plant protection from the drought and other stresses. Annexins have been identified as a component of signal-transduction pathways in many species, such as soybean (Feng et al., 2013) and rice (Jami et al., 2012). In the study performed by Konopka-Postupolska and Clark, overexpression of annexin 1 (AnnAt1) not only allowed plants to retain their growth and productivity potentials under severe drought stress condition but also provided them protection against oxidative stress, however the mechanism of protection mediated by AnnAt1 can be different in different tissues (Konopka-Postupolska and Clark, 2017). Similarly, defense-related proteins such as diseases resistance response protein also showed positive correlation scores with increased root growth of 843-22B (Figure 8A) indicating activation of defense-related pathways.

Two isoforms of phenylalanine ammonia-lyase showed a positive correlation with the tolerant ICTP8203 root phenotypes (Figure 8A) suggesting an activated biosynthesis of antioxidative compounds. Phenylpropanoid compounds not only fulfill various essential functions during plant development but also they act as essential protectants against various biotic and abiotic environmental stresses. Two isoforms of peroxidases also showed a positive correlation with the ICTP8203 root phenotype. Peroxidases are the bifunctional enzymes which not only act on oxidizing agents but also produce ROS (Passardi et al., 2004). In plants, this enzyme is involved in many physiological and developmental processes which include their association with cell elongation processes, and also with reactions that restrict growth.

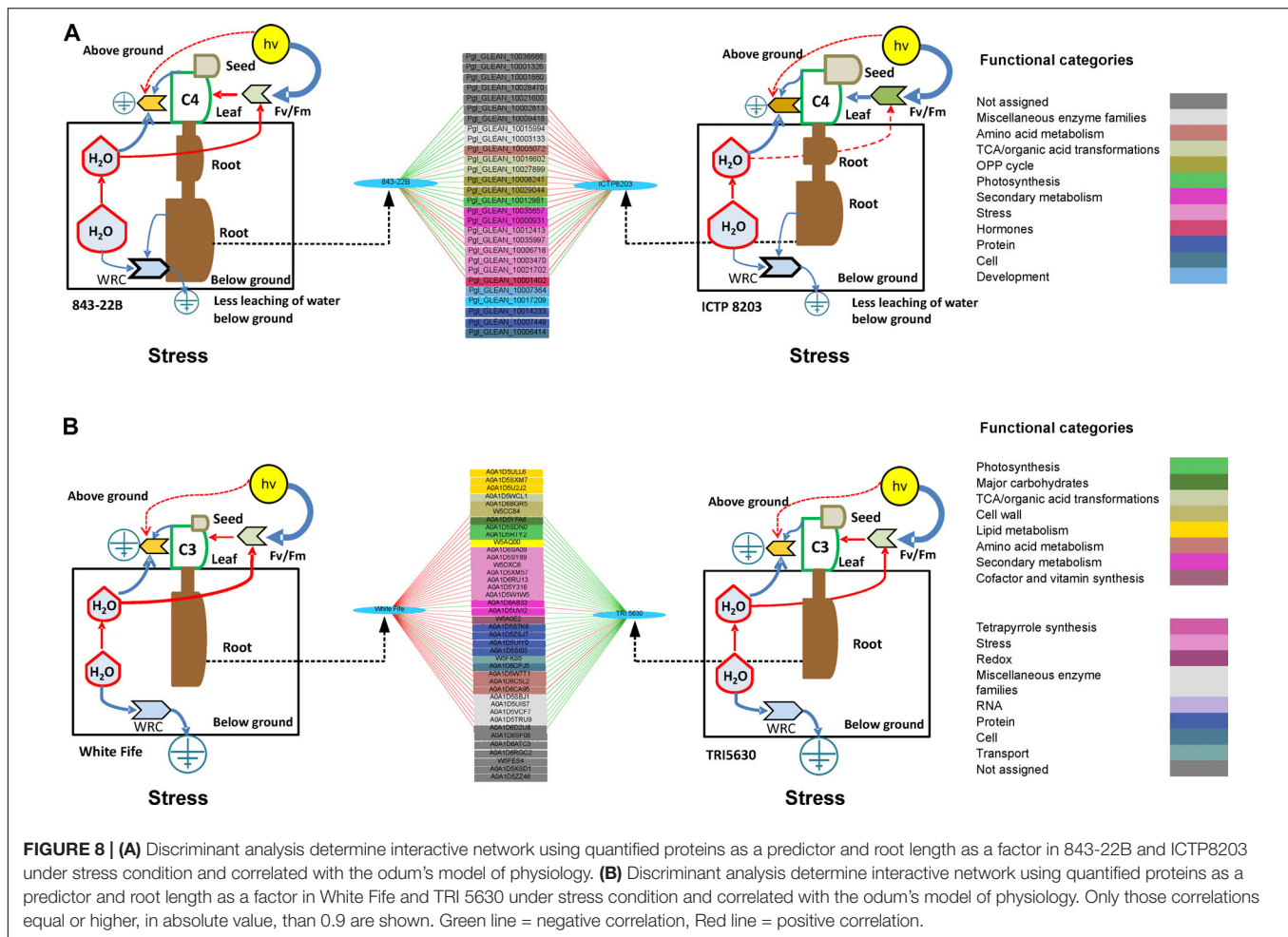


FIGURE 8 | (A) Discriminant analysis determine interactive network using quantified proteins as a predictor and root length as a factor in 843-22B and ICTP8203 under stress condition and correlated with the odum's model of physiology. **(B)** Discriminant analysis determine interactive network using quantified proteins as a predictor and root length as a factor in White Fife and TRI 5630 under stress condition and correlated with the odum's model of physiology. Only those correlations equal or higher, in absolute value, than 0.9 are shown. Green line = negative correlation, Red line = positive correlation.

Reduction in cell wall plasticity leads to the stiffening of the cell wall. Covalently bound cell wall peroxidases play a significant role in this process either by catalyzing the polymerization of the phenolic monomers of lignin or by participating in the formation of cross-bridges between various polysaccharide polymers. Based on the identified protein candidates, it can be concluded that under drought stress, antioxidant activity was enhanced in ICTP8203 compared to 843-22B.

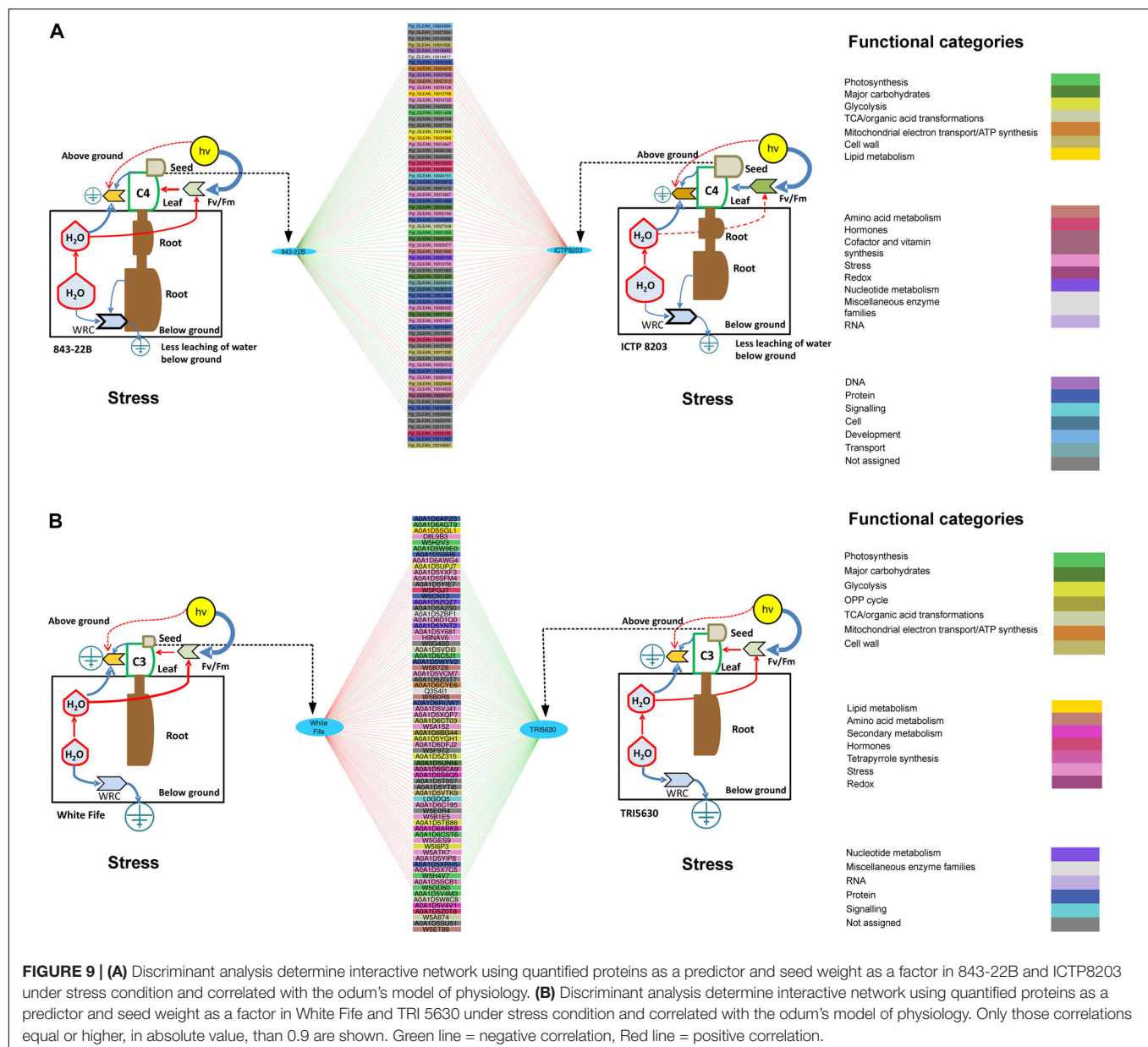
The same approach was applied to the stress proteome of the wheat genotypes and correlated to the root phenotypes (Supplementary Figure S8 and Supplementary Table S16). This analysis revealed rather different patterns compared to pearl millet except peroxidases, which were also identified. In the discriminant analysis, proteins majorly binned in the functional category of lipid metabolism, stress, amino acid metabolism, carbohydrate metabolism and secondary metabolism were determined (Figure 8B). In the physiological analysis of White Fife and TRI 5630 showed increased root length under stress condition compared to controls. Two isoforms of eukaryotic translation initiation factor 3 (eIF3) showed a positive correlation with White Fife (Figure 8B). A study performed by Singh and coworker demonstrated that the overexpression of a gene encoding eIF3g (TaeIF3g;

Triticum aestivum eukaryotic initiation factors), one of the 11 subunits of eukaryotic translation initiation factor 3 (eIF3), showed enhanced tolerance to abiotic stress in yeast and transgenic lines of *Arabidopsis* (Singh et al., 2013). β -glucosidase is actively involved in cell-wall modification, stress defense, phytohormone signaling, and secondary metabolism; it also plays an essential role in the hydrolysis of cellulose by converting cellobiose to glucose. Here in this study, β -glucosidase and β -amylase showed a positive correlation with White Fife, which indicates that carbohydrate metabolism was activated in the roots of White Fife under drought stress condition. In soybean roots, β -glucosidase showed higher accumulation under stress condition (Wang and Komatsu, 2018).

Drought stress has adverse effects on seed production; however, it also largely depends upon the duration of drought stress, the growth stage of the plant and seed filling period. In the present study, genotypes of wheat and pearl millet showed contrasting effects of drought on seed productivity. Based on the physiological analysis, pearl millet genotypes were able to maintain productivity under stress conditions compared to the wheat genotypes, especially ICTP8203. The stress proteome of pearl millet genotypes was correlated with the seed weight (Supplementary Figure S9 and Supplementary

Table S15) and proteins binned in the functional category of stress, carbohydrate metabolism, lipid metabolism, amino acid metabolism, signaling and development were identified (**Figure 9A; Supplementary Table S15**). Two isoforms of late embryogenesis abundant (LEA) protein and lipid transfer protein (LTP) were identified and showed a positive correlation with ICTP8203. These proteins were also differentially abundant in seed proteome of ICTP8203 and not in 843-22B, which confer that these proteins can be used as potential tissue-specific genotype marker for ICTP8203 (**Figure 9A**). Glycogen synthase which is binned in the functional category of carbohydrate metabolism also showed a positive correlation with ICTP8203, and it has been reported that glycogen synthase confers enhanced tolerance against abiotic stresses (Joshi et al., 2018; **Figure 9A**).

Similarly, two isoforms of heat shock proteins were also identified, showing a positive correlation with ICTP8203. Many small heat shock proteins (sHSPs) play a major role in the protection of seeds from desiccation. This also indicates that accumulation of protective proteins might play a significant role in the drought-responsive mechanism of ICTP8203, leading to healthy seeds even under water deficit condition; indeed, the productivity of plants does not get as severely affected as for other compared genotypes under drought stress. Further, the chitinase protein was identified, which showed a positive correlation with 843-22B (**Figure 9A**). In plants, chitinases are induced in defense response against abiotic stress and constitutively expressed in plant organs, such as seeds. A class 1 chitinase has been previously identified as an abundant



protein in the soybean (*Glycine max*) mature seed coat (Gijzen et al., 2001).

SPLS and discriminant analysis of wheat seed proteome lead us to identify proteins binned in the functional category of reactive oxygen species (ROS), photosynthesis, stress, glycolysis, and secondary metabolism (Figure 9B, Supplementary Figure S10, and Supplementary Table S16). Several isoforms of histones were identified, which showed a positive correlation with White Fife. Increasing evidence shows that chromatin organization plays a very important role in the transcriptional reprogramming of stress-responsive gene expression, proper resource allocation to growth vs. stress responses, acclimation, and long-term stress memory (Chinnusamy and Zhu, 2009; Mirouze and Paszkowski, 2011). Several studies have also reported that histones play an important role during the process of grain filling and drought stress leads to the modification of this protein (Kalamajka et al., 2010). Further, two isoforms of superoxide dismutase were identified, which showed a positive correlation with White Fife, indicating that scavenging mechanism was activated under drought stress. Similarly, two isoforms of phosphoglycerate kinase also showed a positive correlation with White Fife. Phosphoglycerate kinase is an enzyme that catalyzes the reversible transfer of a phosphate group from 1, 3-bisphosphoglycerate (1, 3-BPG) to ADP producing 3-phosphoglycerate (3-PG) and ATP (Fermo et al., 2012). These results indicate that energy metabolism was activated under drought stress.

CONCLUSION

The need for drought-tolerant crops is critical, and will surely grow in coming years because of the increase in food demand per capita, ongoing degradation of soil, depletion of water resources, and the accelerating effects of global climate change. Therefore, the development of improved drought-tolerant varieties is the challenge for plant breeders and crop physiologists. One strategy to increase crop productivity is to endow them with C_4 traits. However, this approach needs a better understanding of the molecular and physiological level. Here, we exploited a comparative approach to define molecular and physiological components of C_3 and C_4 plants under drought stress and identified the traits which can be of help in the ongoing engineering process. The results herein presented, reflect the complex mechanism at physiological and proteome level that C_3 wheat and C_4 pearl millet employ in adapting to the drought stress environment, which also testifies toward the plasticity of these plants. An intriguing result is that tolerant lines of pearl millet and wheat seem to override already differences between C_3 - and C_4 -type photosynthesis. These results are subject to future studies. Our results also provide evidence that one of the most substantial advantages for enhanced drought-tolerance is the stay-green trait, and we demonstrate the first proteome signature in pearl millet for this trait. Other important characteristics of drought resistance are related to root morphology, efficient photosynthetic machinery and wax biosynthetic enzymes.

We suggest several strategies for engineering enhanced tolerance in the crop plants under drought stress: (1) Identification and mapping of quantitative trait loci (QTL) for root length in cereal crops as it is an important trait for survival under drought stress. (2) Enrichment and expansion of stay-green protein signatures and wax biosynthetic genes. (3) Further investigation of reactive oxygen species (ROS) in roots as they can act as efficient signaling molecules which can enhance root to shoot crosstalk under drought stress. This study also provides information on yield-associated traits, tissue and genotype specific marker that can be exploited for marker-assisted breeding for improving drought tolerance in crop plants. This knowledge is very important because of the large discrepancy between gene expression level and protein activities, which are dynamically modified by actual field conditions in a strongly fluctuating climate. To our knowledge, it is the first report on a comparative physiological and proteomic analysis of wheat and pearl millet in response to drought stress, thus, serving as a large-scale reference study for future investigations.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

PC, AG, and WW conceived and designed the experiments and wrote the manuscript. AG, PC, and GB performed the experiments. PC, AG, GB, LV, ŽR, MB, PB, SJ, WL, XS, KG, RKV, and WW analyzed the data. All authors contributed to editing and agreed on the final version.

FUNDING

We thank Austrian Science Fund (FWF, DerWissenschaftsfonds), Grant agreement number W1257 for the financial support of AG. LV was supported by Spanish Ministry of Science and Universities (RYC-2015-17871). ŽR and KG were supported by the Slovenian Research Agency (grants P4-0165, J4-1777, and Z7-1888). RKV thanks CGIAR Research Program on Grain Legumes and Dryland Cereals for Funding Research in Pearl Millet at Center of Excellence in Genomics & Systems Biology (CEGSB), ICRIAT, Patancheru, Hyderabad, India.

ACKNOWLEDGMENTS

We are thankful to Ms. Annapurna Chitikineni, Senior Manager CEGSB, ICRIAT, Patancheru, Hyderabad, India, Dr. Michele Grieco from IPK Gatersleben, Germany, and Dr. Markus Teige from Max Perutz Labs, Vienna, Austria for arranging seeds of pearl millet and wheat, respectively. We would also like

to thank the gardeners Andreas Schröfl and Thomas Joch for excellent plant cultivation in the glasshouse facility of Molecular Systems Biology Lab (MOSYS), Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.600278/full#supplementary-material>

Supplementary Figure 1 | (A) Orthogonal partial least squares discriminant analysis (OPLS-DA) was performed considering all the factors and variables of physiological parameters under control and stress condition in pearl millet and wheat genotypes. **(B)** Principal component analysis (PCA) of all the physiological parameters under control and stress in pearl millet and wheat genotypes.

Supplementary Figure 2 | Venn diagrams represent total proteome identification in root, leaf and seed tissues of pearl millet and wheat genotypes under control and stress condition.

Supplementary Figure 3 | Principal component analysis (PCA) of pearl millet genotypes. **(A)** PCA considering all the tissues of pearl millet tolerant and sensitive genotypes (84322B and ICTP8203) under control and stress. **(B)** PCA of ICTP8203 under control and drought stress in roots, seeds, and leaves. **(C)** Loading graphs of average for 100 positive and 100 negative loadings. **(D)** PCA of 843-22B under control and drought stress in roots, seeds, and leaves; **(E)** Loading graphs of average for 100 positive and 100 negative loadings.

Supplementary Figure 4 | Principal component analysis (PCA) of wheat genotypes. **(A)** PCA considering all the tissues of wheat tolerant and sensitive genotypes (White Fife and TRI 5630) under control and stress. **(B)** PCA of TRI 5630 under control and drought stress in roots, seeds, and leaves. **(C)** Loading graphs of average for 100 positive and 100 negative loadings. **(D)** PCA of White Fife under control and drought stress in roots, seeds, and leaves. **(E)** Loading graphs of average for 100 positive and 100 negative loadings.

Supplementary Figure 5 | Root length measurements: Box plots represent root length of the plants under control and stress condition in pearl millet and wheat genotypes; Root length pictures of pearl millet and wheat plants under control and stress condition.

Supplementary Figure 6 | Regulation of C₄ photosynthesis pathway proteins in pearl millet (843-22B and ICTP8203) and wheat (White Fife and TRI 5630) genotypes under control and stress condition.

Supplementary Figure 7 | sPLS-based network constructed using quantified proteins as a predictor and root length as a factor for pearl millet genotypes (843-22B and ICTP8203) under stress condition. The interacting network was correlated with the odum's model of physiology. Only those correlations equal or higher, in absolute value, than 0.9 are shown. Green line = negative correlation, Red line = positive correlation.

Supplementary Figure 8 | sPLS-based network constructed using quantified proteins as a predictor and root length as factor for wheat genotypes (White Fife and TRI5630) under stress condition. The interacting network was correlated with the odum's model of physiology. Only those correlations equal or higher, in absolute value, than 0.9 are shown. Green line = negative correlation, Red line = positive correlation.

Supplementary Figure 9 | sPLS-based network constructed using quantified proteins as a predictor and seed weight as factor for pearl millet genotypes (84322B and ICTP8203) under stress condition. Interacting network was correlated with the odum's model of physiology. Only those correlations equal or higher, in absolute value, than 0.9 are shown. Green line = negative correlation, Red line = positive correlation.

Supplementary Figure 10 | sPLS-based network constructed using quantified proteins as a predictor and seed weight as factor for wheat genotypes (White Fife and TRI 5630) under stress condition. The interacting network was correlated with the odum's model of physiology. Only those correlations equal or higher, in absolute value, than 0.9 are shown. Green line = negative correlation, Red line = positive correlation.

Supplementary Table 1 | Details of the sample code for the raw files measured and submitted in PRIDE repository for Wheat and Pearl millet genotypes.

Supplementary Table 2 | Physiological parameters under control and drought stress in pearl millet and wheat genotypes.

Supplementary Table 3 | Qualitative and quantitative identification of proteins in pearl millet (843-22B and ICTP8203) genotypes.

Supplementary Table 4 | Qualitative and quantitative identification of proteins in wheat (White Fife and TRI5630) genotypes.

Supplementary Table 5 | Two-way ANOVA analysis of pearl millet (843-22B; ICTP8203) and wheat (White Fife; TRI5630) genotypes.

Supplementary Table 6 | Differentially expressed proteins in root, leaf, seed tissues of pearl millet sensitive (843-22B) and tolerant (ICTP8203) genotypes.

Supplementary Table 7 | Differentially expressed proteins in root, leaf, seed tissues of wheat sensitive (White Fife) and tolerant (TRI 5630) genotypes.

Supplementary Table 8 | Cluster analysis of pearl millet (843-22B and ICTP8203) genotypes using Kmeans in Matlab.

Supplementary Table 9 | Cluster analysis of wheat (White Fife and TRI5630) genotypes using Kmeans in Matlab.

Supplementary Table 10 | Functional categorization of the identified proteins in pearl millet (843-22B and ICTP8203) genotypes.

Supplementary Table 11 | Functional categorization of the identified proteins in wheat (White Fife and TRI5630) genotypes.

Supplementary Table 12 | PCA loadings of pearl millet genotypes (843-22B and ICTP8203).

Supplementary Table 13 | PCA loadings of wheat genotypes (White Fife and TRI5630).

Supplementary Table 14 | Proteome regulation underlying 'Stay-green' and 'Senescence' trait in pearl millet genotypes (843-22B and ICTP8203) and wheat genotypes (White Fife and TRI5630).

Supplementary Table 15 | sPLS and discriminant network analysis, pearl millet (843-22B and ICTP8203) stress proteins integrated with root length and seed weight and color-code-bins used for the functional category for network analysis.

Supplementary Table 16 | sPLS and discriminant network analysis, wheat stress (White Fife and TRI5630) proteins integrated with root length and seed weight.

REFERENCES

- Ahmed, K., Shabbir, G., Ahmed, M., and Shah, K. N. (2020). Phenotyping for drought resistance in bread wheat using physiological and biochemical traits. *Sci. Total Environ.* 729:139082. doi: 10.1016/j.scitotenv.2020.139082
- Alexandratos, N., and Bruinsma, J. (2012). *World Agriculture: Towards 2015/2030 an FAO Perspective*. Rome: FAO.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Anjum, S. A., Xie, X. Y., Wang, L. C., Saleem, M. F., Man, C., and Lei, W. (2011). Morphological, physiological and biochemical responses of plants to drought stress. *Afr. J. Agric. Res.* 6, 2026–2032.

- Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191.
- Ashraf, M. Y., Azmi, A. R., Khan, A. H., and Ala, S. A. (1994). Effect of water-stress on total phenols, peroxidase-activity and chlorophyll content in wheat (*Triticum-Aestivum* L.). *Acta Physiol. Plant.* 16, 185–191.
- Baker, N. R., and Rosenqvist, E. (2004). Applications of chlorophyll fluorescence can improve crop production strategies: an examination of future possibilities. *J. Exp. Bot.* 55, 1607–1621. doi: 10.1093/jxb/erh196
- Bateman, A., Martin, M. J., O'donovan, C., Magrane, M., Alpi, E., Antunes, R., et al. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45, D158–D169.
- Begg, J. E. (1965). Growth and development of a crop of bulrush millet (*Pennisetum typhoides* S + H). *J. Agric. Sci.* 65, 341–349. doi: 10.1017/s0021859600048905
- Bidinger, F. R., Mahalakshmi, V., and Rao, G. D. P. (1987). Assessment of drought resistance in pearl-millet [*Pennisetum americanum* (L.) Leeke]. 1. Factors affecting yields under stress. *Aust. J. Agric. Res.* 38, 37–48. doi: 10.1071/ar9870037
- Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R., et al. (2007). Characterization of Arabidopsis ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion. *Plant J.* 52, 485–498. doi: 10.1111/j.1365-3113x.2007.03252.x
- Borrell, A. K., and Hammer, G. L. (2000). Nitrogen dynamics and the physiological basis of stay-green in sorghum. *Crop Sci.* 40, 1295–1307. doi: 10.2135/cropsci2000.4051295x
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Buckley, T. N. (2019). How do stomata respond to water status? *New Phytol.* 224, 21–36. doi: 10.1111/nph.15899
- Campos, H., Cooper, A., Habben, J. E., Edmeades, G. O., and Schussler, J. R. (2004). Improving drought tolerance in maize: a view from industry. *Field Crops Res.* 90, 19–34. doi: 10.1016/j.fcr.2004.07.003
- Chaturvedi, P., Doerfler, H., Jegadeesan, S., Ghatak, A., Pressman, E., Castillejo, M. A., et al. (2015). Heat-treatment-responsive proteins in different developmental stages of tomato pollen detected by targeted mass accuracy precursor alignment (tMAPA). *J. Proteome Res.* 14, 4463–4471. doi: 10.1021/pr501240n
- Chaturvedi, P., Ischebeck, T., Egelhofer, V., Lichtscheid, I., and Weckwerth, W. (2013). Cell-specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen provides evidence for developmental priming. *J. Proteome Res.* 12, 4892–4903. doi: 10.1021/pr400197p
- Cheng, C. Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J.* 89, 789–804. doi: 10.1111/tpj.13415
- Chia, T., Thorneycroft, D., Chapple, A., Messerli, G., Chen, J., Zeeman, S. C., et al. (2004). A cytosolic glucosyltransferase is required for conversion of starch to sucrose in *Arabidopsis* leaves at night. *Plant J.* 37, 853–863. doi: 10.1111/j.1365-3113x.2003.02012.x
- Chinnusamy, V., and Zhu, J. K. (2009). Epigenetic regulation of stress responses in plants. *Curr. Opin. Plant Biol.* 12, 133–139. doi: 10.1016/j.pbi.2008.12.006
- Chmielewska, K., Rodziewicz, P., Swarczewicz, B., Sawikowska, A., Krajewski, P., Marczak, L., et al. (2016). Analysis of drought-induced proteomic and metabolomic changes in barley (*Hordeum vulgare* L.) leaves and roots unravels some aspects of biochemical mechanisms involved in drought tolerance. *Front. Plant Sci.* 7:1108. doi: 10.3389/fpls.2016.01108
- Comas, L. H., Becker, S. R., Cruz, V. V., Byrne, P. F., and Dierig, D. A. (2013). Root traits contributing to plant productivity under drought. *Front. Plant Sci.* 4:442. doi: 10.3389/fpls.2013.00442
- de Dorlodot, S., Forster, B., Pages, L., Price, A., Tuberosa, R., and Draye, X. (2007). Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends Plant Sci.* 12, 474–481. doi: 10.1016/j.tplants.2007.08.012
- Debieu, M., Sine, B., Passot, S., Grondin, A., Akata, E., Gangashetty, P., et al. (2018). Response to early drought stress and identification of QTLs controlling biomass production under drought in pearl millet. *PLoS One* 13:e0201635. doi: 10.1371/journal.pone.0201635
- Ding, Z. S., Sun, X. F., Huang, S. H., Zhou, B. Y., and Zhao, M. (2015). Response of photosynthesis to short-term drought stress in rice seedlings overexpressing C-4 phosphoenolpyruvate carboxylase from maize and millet. *Photosynthetica* 53, 481–488. doi: 10.1007/s11099-015-0126-1
- Edwards, G., and Walker, D. (1983). *C₃ C₄: Mechanisms, and Cellular and Environmental Regulation, of Photosynthesis*. Oxford: Blackwell Scientific Publications.
- Escandon, M., Valledor, L., Pascual, J., Pinto, G., Canal, M. J., and Meijon, M. (2017). System-wide analysis of short-term response to high temperature in *Pinus radiata*. *J. Exp. Bot.* 68, 3629–3641. doi: 10.1093/jxb/erx198
- Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., and Basra, S. M. A. (2009). Plant drought stress: effects, mechanisms and management. *Agron. Sustain. Dev.* 29, 185–212. doi: 10.1051/agro:2008021
- Feng, Y. M., Wei, X. K., Liao, W. X., Huang, L. H., Zhang, H., Liang, S. C., et al. (2013). Molecular analysis of the annexin gene family in soybean. *Biol. Plant.* 57, 655–662. doi: 10.1007/s10535-013-0334-0
- Fermo, E., Bianchi, P., Chiarelli, L. R., Maggi, M., Mandara, G. M. L., Vercellati, C., et al. (2012). A new variant of phosphoglycerate kinase deficiency (p.I371K) with multiple tissue involvement: molecular and functional characterization. *Mol. Genet. Metab.* 106, 455–461. doi: 10.1016/j.ymgme.2012.05.015
- Fischer, R. A., and Wood, J. T. (1979). Drought resistance in spring wheat cultivars. 3. Yield associations with morpho-physiological traits. *Aust. J. Agric. Res.* 30, 1001–1020. doi: 10.1071/ar9791001
- Ford, K. L., Cassin, A., and Bacic, A. (2011). Quantitative proteomic analysis of wheat cultivars with differing drought stress tolerance. *Front. Plant Sci.* 2:44. doi: 10.3389/fpls.2011.00044
- Fu, L. M., Niu, B. F., Zhu, Z. W., Wu, S. T., and Li, W. Z. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Ghannoum, O. (2009). C-4 photosynthesis and water stress. *Ann. Bot.* 103, 635–644. doi: 10.1093/aob/mcn093
- Ghatak, A., Chaturvedi, P., Nagler, M., Roustan, V., Lyon, D., Bachmann, G., et al. (2016). Comprehensive tissue-specific proteome analysis of drought stress responses in *Pennisetum glaucum* (L.) R. Br. (Pearl millet). *J. Proteomics* 143, 122–135. doi: 10.1016/j.jprot.2016.02.032
- Ghatak, A., Chaturvedi, P., Paul, P., Agrawal, G. K., Rakwal, R., Kim, S. T., et al. (2017a). Proteomics survey of Solanaceae family: current status and challenges ahead. *J. Proteomics* 169, 41–57. doi: 10.1016/j.jprot.2017.05.016
- Ghatak, A., Chaturvedi, P., and Weckwerth, W. (2017b). Cereal crop proteomics: systemic analysis of crop drought stress responses towards marker-assisted selection breeding. *Front. Plant Sci.* 8:757. doi: 10.3389/fpls.2017.00757
- Ghatak, A., Chaturvedi, P., and Weckwerth, W. (2018). Metabolomics in plant stress physiology. *Adv. Biochem. Eng. Biotechnol.* 164, 187–236. doi: 10.1007/10_2017_55
- Gijzen, M., Kuflu, K., Qutob, D., and Chernys, J. T. (2001). A class I chitinase from soybean seed coat. *J. Exp. Bot.* 52, 2283–2289. doi: 10.1093/jxb/52.365.2283
- Hanjra, M. A., and Qureshi, M. E. (2010). Global water crisis and future food security in an era of climate change. *Food Policy* 35, 365–377. doi: 10.1016/j.foodpol.2010.05.006
- Harris, K., Subudhi, P. K., Borrell, A., Jordan, D., Rosenow, D., Nguyen, H., et al. (2007). Sorghum stay-green QTL individually reduce postflowering drought-induced leaf senescence. *J. Exp. Bot.* 58, 327–338. doi: 10.1093/jxb/erl225
- Hatch, M. D. (1987). C-4 photosynthesis - a unique blend of modified biochemistry, anatomy and ultrastructure. *Biochim. Biophys. Acta* 895, 81–106. doi: 10.1016/s0304-4173(87)80009-5
- Hoehenwarter, W., Van Dongen, J. T., Wienkoop, S., Steinfath, M., Hummel, J., Erban, A., et al. (2008). A rapid approach for phenotype-screening and database independent detection of cSNP/protein polymorphism using mass accuracy precursor alignment. *Proteomics* 8, 4214–4225. doi: 10.1002/pmic.200701047
- Holmgren, M., Gomez-Aparicio, L., Quero, J. L., and Valladares, F. (2012). Non-linear effects of drought under shade: reconciling physiological and ecological models in plant communities. *Oecologia* 169, 293–305. doi: 10.1007/s00442-011-2196-5
- Huber, S. C., Bachmann, M., and Huber, J. L. (1996). Post-translational regulation of nitrate reductase activity: a role for Ca²⁺ and 14-3-3 proteins. *Trends Plant Sci.* 1, 432–438. doi: 10.1016/s1360-1385(96)10046-7
- Ignatova, L., Rudenko, N., Zhurikova, E., Borisova-Mubarakshina, M., and Ivanov, B. (2019). Carbonic anhydrases in photosynthesizing cells of C₃ higher plants. *Metabolites* 9:73. doi: 10.3390/metabo9040073
- Izanloo, A., Condon, A. G., Langridge, P., Tester, M., and Schnurbusch, T. (2008). Different mechanisms of adaptation to cyclic water stress in two South Australian bread wheat cultivars. *J. Exp. Bot.* 59, 3327–3346. doi: 10.1093/jxb/ern199

- Jami, S. K., Clark, G. B., Ayele, B. T., Ashe, P., and Kirti, P. B. (2012). Genome-wide comparative analysis of annexin superfamily in plants. *PLoS One* 7:e47801. doi: 10.1371/journal.pone.0047801
- Jenner, H. L., Winning, B. M., Millar, A. H., Tomlinson, K. L., Leaver, C. J., and Hill, S. A. (2001). NAD malic enzyme and the control of carbohydrate metabolism in potato tubers. *Plant Physiol.* 126, 1139–1149. doi: 10.1104/pp.126.3.1139
- Jiang, G. H., He, Y. Q., Xu, C. G., Li, X. H., and Zhang, Q. (2004). The genetic basis of stay-green in rice analyzed in a population of doubled haploid lines derived from an indica by japonica cross. *Theor. Appl. Genet.* 108, 688–698. doi: 10.1007/s00122-003-1465-z
- Jiao, D. M., Huang, X. Q., Li, X., Chi, W., Kuang, T. Y., Zhang, Q. D., et al. (2002). Photosynthetic characteristics and tolerance to photo-oxidation of transgenic rice expressing C(4) photosynthesis enzymes. *Photosynth. Res.* 72, 85–93.
- Jones, H. G. (1974). Assessment of stomatal control of plant water status. *New Phytol.* 73, 851–859. doi: 10.1111/j.1469-8137.1974.tb01314.x
- Jones, H. G. (2007). Monitoring plant and soil water status: established and novel methods revisited and their relevance to studies of drought tolerance. *J. Exp. Bot.* 58, 119–130. doi: 10.1093/jxb/erl118
- Joshi, R., Singla-Pareek, S. L., and Pareek, A. (2018). Engineering abiotic stress response in plants for biomass production. *J. Biol. Chem.* 293, 5035–5043. doi: 10.1074/jbc.tml117.000232
- Kalamajka, R., Finnie, C., and Grasser, K. D. (2010). Onset of grain filling is associated with a change in properties of linker histone variants in maize kernels. *Planta* 231, 1127–1135. doi: 10.1007/s00425-010-1119-8
- Kamal, N. M., Gorafi, Y. S. A., Abdelrahman, M., Abdellatif, E., and Tsujimoto, H. (2019). Stay-green trait: a prospective approach for yield potential, and drought and heat stress adaptation in globally important cereals. *Int. J. Mol. Sci.* 20:5837. doi: 10.3390/ijms20235837
- Katam, R., Sakata, K., Suravajhala, P., Pechan, T., Kambiranda, D. M., Naik, K. S., et al. (2016). Comparative leaf proteomics of drought-tolerant and -susceptible peanut in response to water stress. *J. Proteom.* 143, 209–226. doi: 10.1016/j.jprot.2016.05.031
- Kawahara, Y., De La Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4. doi: 10.1186/1939-8433-6-4
- Komatsu, S., Kamal, A. H. M., and Hossain, Z. (2014). Wheat proteomics: proteome modulation and abiotic stress acclimation. *Front. Plant Sci.* 5:684. doi: 10.3389/fpls.2014.00684
- Konopka-Postupolska, D., and Clark, G. (2017). Annexins as overlooked regulators of membrane trafficking in plant cells. *Int. J. Mol. Sci.* 18:863. doi: 10.3390/ijms18040863
- Lamalakshmi Devi, E., Kumar, S., Singh, B. T., Sharma, S. K., Beemrote, A., Premabati Devi, C., et al. (2017). *Adaptation Strategies and Defence Mechanisms of Plants During Environmental Stress*. Cham: Springer International Publishing.
- Lee, S. B., and Suh, M. C. (2015). Cuticular wax biosynthesis is up-regulated by the MYB94 transcription factor in Arabidopsis. *Plant Cell Physiol.* 56, 48–60. doi: 10.1093/pcp/pcu142
- Leegood, R. C. (2002). C(4) photosynthesis: principles of CO(2) concentration and prospects for its introduction into C(3) plants. *J. Exp. Bot.* 53, 581–590. doi: 10.1093/jexbot/53.369.581
- Leishman, M. R., and Westoby, M. (1994). The role of seed size in seedling establishment in dry soil-conditions - experimental-evidence from semiarid species. *J. Ecol.* 82, 249–258. doi: 10.2307/2261293
- Li, L., Stoeckert, C. J., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503
- Liu, F., and Stutzel, H. (2002). Leaf expansion, stomatal conductance, and transpiration of vegetable amaranth (*Amaranthus* sp.) in response to soil drying. *J. Am. Soc. Hortic. Sci.* 127, 878–883. doi: 10.21273/jashs.127.5.878
- Liu, F. L., Andersen, M. N., Jacobsen, S. E., and Jensen, C. R. (2005). Stomatal control and water use efficiency of soybean (*Glycine max* L. Merr.) during progressive soil drying. *Environ. Exp. Bot.* 54, 33–40. doi: 10.1016/j.envexpbot.2004.05.002
- Loggini, B., Scartazza, A., Brugnoli, E., and Navari-Izzo, F. (1999). Antioxidative defense system, pigment composition, and photosynthetic efficiency in two wheat cultivars subjected to drought. *Plant Physiol.* 119, 1091–1099. doi: 10.1104/pp.119.3.1091
- Lu, C. M., and Zhang, J. H. (1998). Changes in photosystem II function during senescence of wheat leaves. *Physiol. Plant.* 104, 239–247. doi: 10.1034/j.1399-3054.1998.1040212.x
- Lu, Q. T., Lu, C. M., Zhang, J. H., and Kuang, T. Y. (2002). Photosynthesis and chlorophyll a fluorescence during flag leaf senescence of field-grown wheat plants. *J. Plant Physiol.* 159, 1173–1178. doi: 10.1078/0176-1617-00727
- Ludlow, M. M., and Muchow, R. C. (1990). A critical-evaluation of traits for improving crop yields in water-limited environments. *Adv. Agron.* 43, 107–153. doi: 10.1016/s0065-2113(08)60477-0
- Lunt, T., Jones, A. W., Mulhern, W. S., Lezaks, D. P. M., and Jahn, M. M. (2016). Vulnerabilities to agricultural production shocks: an extreme, plausible scenario for assessment of risk for the insurance sector. *Clim. Risk Manag.* 13, 1–9. doi: 10.1016/j.crm.2016.05.001
- Mabhaudhi, T., Chimonyo, V. G. P., Hlahla, S., Massawe, F., Mayes, S., Nhamo, L., et al. (2019). Prospects of orphan crops in climate change. *Planta* 250, 695–708. doi: 10.1007/s00425-019-03129-y
- Michaelletti, A., Naghavi, M. R., Toorchi, M., Zolla, L., and Rinalducci, S. (2018). Metabolomics and proteomics reveal drought-stress responses of leaf tissues from spring-wheat. *Sci. Rep.* 8:5710.
- Mirouze, M., and Paszkowski, J. (2011). Epigenetic contribution to stress adaptation in plants. *Curr. Opin. Plant Biol.* 14, 267–274. doi: 10.1016/j.pbi.2011.03.004
- Munne-Bosch, S., and Alegre, L. (2000). Changes in carotenoids, tocopherols and diterpenes during drought and recovery, and the biological significance of chlorophyll loss in *Rosmarinus officinalis* plants. *Planta* 210, 925–931. doi: 10.1007/s004250050699
- Munne-Bosch, S., and Alegre, L. (2004). Die and let live: leaf senescence contributes to plant survival under drought stress. *Funct. Plant Biol.* 31, 203–216. doi: 10.1071/fp03236
- Nagahatenna, D. S. K., Langridge, P., and Whitford, R. (2015). Tetrapyrrole-based drought stress signalling. *Plant Biotechnol. J.* 13, 447–459. doi: 10.1111/pbi.12356
- Nagler, M., Nagele, T., Gilli, C., Fragner, L., Korte, A., Platzer, A., et al. (2018). Eco-metabolomics and metabolic modeling: making the leap from model systems in the lab to native populations in the field. *Front. Plant Sci.* 9:1556. doi: 10.3389/fpls.2018.01556
- Odum, H. T. (1994). Ecological engineering - the necessary use of ecological self-design. *Ecol. Eng.* 3, 115–118.
- Panikashvili, D., Savaldi-Goldstein, S., Mandel, T., Yifhar, T., Franke, R. B., Hofer, R., et al. (2007). The Arabidopsis *DESPERADO/AtWBC11* transporter is required for cutin and wax secretion. *Plant Physiol.* 145, 1345–1360. doi: 10.1104/pp.107.105676
- Paoletti, A. C., Parmely, T. J., Tomomori-Sato, C., Sato, S., Zhu, D. X., Conaway, R. C., et al. (2006). Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18928–18933. doi: 10.1073/pnas.0606379103
- Passardi, F., Penel, C., and Dunand, C. (2004). Performing the paradoxical: how plant peroxidases modify the cell wall. *Trends Plant Sci.* 9, 534–540. doi: 10.1016/j.tplants.2004.09.002
- Pighin, J. A., Zheng, H., Balakshin, L. J., Goodman, I. P., Western, T. L., Jetter, R., et al. (2004). Plant cuticular lipid export requires an ABC transporter. *Science* 306, 702–704. doi: 10.1126/science.1102331
- Pinheiro, C., and Chaves, M. M. (2011). Photosynthesis and drought: can we make metabolic connections from available data? *J. Exp. Bot.* 62, 869–882. doi: 10.1093/jxb/erq340
- Pou, A., Flexas, J., Alsina, M. D., Bota, J., Carambula, C., De Herralde, F., et al. (2008). Adjustments of water use efficiency by stomatal regulation during drought and recovery in the drought-adapted Vitis hybrid Richter-110 (*V. berlandieri* x *V. rupestris*). *Physiol. Plant.* 134, 313–323. doi: 10.1111/j.1399-3054.2008.01138.x
- Ramsak, Z., Baebler, S., Rotter, A., Korbar, M., Mozetic, I., Usadel, B., et al. (2014). GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. *Nucleic Acids Res.* 42, D1167–D1175.

- Rangan, P., Furtado, A., and Henry, R. J. (2016). New evidence for grain specific C₄ photosynthesis in wheat. *Sci. Rep.* 6:31721.
- Riccardi, F., Gazeau, P., Jacquemot, M. P., Vincent, D., and Zivy, M. (2004). Deciphering genetic variations of proteome responses to water deficit in maize leaves. *Plant Physiol. Biochem.* 42, 1003–1011. doi: 10.1016/j.plaphy.2004.09.009
- Rivero, R. M., Kojima, M., Gepstein, A., Sakakibara, H., Mittler, R., Gepstein, S., et al. (2007). Delayed leaf senescence induces extreme drought tolerance in a flowering plant. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19631–19636. doi: 10.1073/pnas.0709453104
- Rodziewicz, P., Chmielewska, K., Sawikowska, A., Marczak, L., Luczak, M., Bednarek, P., et al. (2019). Identification of drought responsive proteins and related proteomic QTLs in barley. *J. Exp. Bot.* 70, 2823–2837. doi: 10.1093/jxb/erz075
- Samuels, L., Kunst, L., and Jetter, R. (2008). Sealing plant surfaces: cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol.* 59, 683–707. doi: 10.1146/annurev.arplant.59.103006.093219
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., et al. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/nature11119
- Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., et al. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* 24, 447–454. doi: 10.1038/nbt1192
- Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., Taunk, J., et al. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS One* 10:e0122165. doi: 10.1371/journal.pone.0122165
- Seo, P. J., Xiang, F. N., Qiao, M., Park, J. Y., Lee, Y. N., Kim, S. G., et al. (2009). The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in *Arabidopsis*. *Plant Physiol.* 151, 275–289. doi: 10.1104/pp.109.144220
- Shi, H. R., Wang, B., Yang, P. J., Li, Y. B., and Miao, F. (2016). Differences in sugar accumulation and mobilization between sequential and non-sequential senescence wheat cultivars under natural and drought conditions. *PLoS One* 11:e0166155. doi: 10.1371/journal.pone.0166155
- Shi, J. H., Yi, K. K., Liu, Y., Xie, L., Zhou, Z. J., Chen, Y., et al. (2015). Phosphoenolpyruvate carboxylase in *Arabidopsis* leaves plays a crucial role in carbon and nitrogen metabolism. *Plant Physiol.* 167, 671–681. doi: 10.1104/pp.114.254474
- Singh, B., Chauhan, H., Khurana, J. P., Khurana, P., and Singh, P. (2013). Evidence for the role of wheat eukaryotic translation initiation factor 3 subunit g (TaeIF3g) in abiotic stress tolerance. *Gene* 532, 177–185. doi: 10.1016/j.gene.2013.09.078
- Sun, X., and Weckwerth, W. (2012). COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8, 81–93. doi: 10.1007/s11306-012-0399-3
- Taniguchi, M., Kobe, A., Kato, M., and Sugiyama, T. (1995). Aspartate aminotransferase isozymes in *Panicum miliaceum* L., an NAD-malic enzyme-type C₄ plant: comparison of enzymatic properties primary structures, and expression patterns. *Arch. Biochem. Biophys.* 318, 295–306. doi: 10.1006/abbi.1995.1233
- Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939. doi: 10.1111/j.1365-3113x.2004.02016.x
- Thomas, H., and Howarth, C. J. (2000). Five ways to stay green. *J. Exp. Bot.* 51, 329–337. doi: 10.1093/jexbot/51.suppl_1.329
- Thomas, H., and Ougham, H. (2014). The stay-green trait. *J. Exp. Bot.* 65, 3889–3900. doi: 10.1093/jxb/eru037
- Tuberosa, R., and Salvi, S. (2006). Genomics-based approaches to improve drought tolerance of crops. *Trends Plant Sci.* 11, 405–412. doi: 10.1016/j.tplants.2006.06.003
- Tulloch, A. P. (1973). Composition of leaf surface waxes of *Triticum*-species - variation with age and tissue. *Phytochemistry* 12, 2225–2232. doi: 10.1016/0031-9422(73)85124-6
- Valledor, L., Recuenco-Munoz, L., Egelhofer, V., Wienkoop, S., and Weckwerth, W. (2012). The different proteomes of *Chlamydomonas reinhardtii*. *J. Proteomics* 75, 5883–5887. doi: 10.1016/j.jprot.2012.07.045
- Valledor, L., and Weckwerth, W. (2014). An improved detergent-compatible gel-fractionation LC-LTQ-orbitrap-MS workflow for plant and microbial proteomics. *Plant Proteomics Methods Protoc.* 1072, 347–358. doi: 10.1007/978-1-62703-631-3_25
- Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y., et al. (2018). PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res.* 46, D1190–D1196. doi: 10.1093/nar/gkx1002
- Varshney, R. K., Shi, C. C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976.
- von Caemmerer, S., Quick, W. P., and Furbank, R. T. (2012). The development of C-4 rice: current progress and future challenges. *Science* 336, 1671–1672. doi: 10.1126/science.1220177
- Walulu, R. S., Rosenow, D. T., Wester, D. B., and Nguyen, H. T. (1994). Inheritance of the stay green trait in sorghum. *Crop Sci.* 34, 970–972. doi: 10.2135/cropsci1994.0011183X003400040026x
- Wang, L., Czedik-Eysenberg, A., Mertz, R. A., Si, Y. Q., Tohge, T., Nunes-Nesi, A., et al. (2014). Comparative analyses of C-4 and C-3 photosynthesis in developing leaves of maize and rice. *Nat. Biotechnol.* 32, 1158–1165.
- Wang, X., and Komatsu, S. (2018). Proteomic approaches to uncover the flooding and drought stress response mechanisms in soybean. *J. Proteomics* 172, 201–215. doi: 10.1016/j.jprot.2017.11.006
- Weckwerth, W. (2008). Integration of metabolomics and proteomics in molecular plant physiology—coping with the complexity by data-dimensionality reduction. *Physiol. Plant.* 132, 176–189. doi: 10.1111/j.1399-3054.2007.01011.x
- Weckwerth, W. (2019). Toward a unification of system-theoretical principles in biology and ecology—the stochastic Lyapunov matrix equation and its inverse application. *Front. Appl. Math. Stat.* 5:29. doi: 10.3389/fams.2019.00029
- Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P., and Varshney, R. K. (2020). PANOMICS meets germplasm. *Plant Biotechnol. J.* 18, 1507–1525. doi: 10.1111/pbi.13372
- Weckwerth, W., and Morgenthal, K. (2005). Metabolomics: from pattern recognition to biological interpretation. *Drug Discov. Today* 10, 1551–1558. doi: 10.1016/s1359-6446(05)03609-3
- Weissmann, S., and Brutnell, T. P. (2012). Engineering C-4 photosynthetic regulatory networks. *Curr. Opin. Biotechnol.* 23, 298–304. doi: 10.1016/j.copbio.2011.12.018
- Westhoff, P., and Gowik, U. (2004). Evolution of C₄ phosphoenolpyruvate carboxylase. Genes and proteins: a case study with the genus *Flaveria*. *Ann. Bot.* 93, 13–23. doi: 10.1093/aob/mch003
- Wilczynski, G., Kulma, A., and Szopa, J. (1998). The expression of 14-3-3 isoforms in potato is developmentally regulated. *J. Plant Physiol.* 153, 118–126. doi: 10.1016/s0176-1617(98)80054-0
- Xu, W., Rosenow, D. T., and Nguyen, H. T. (2000). Stay green trait in grain sorghum: relationship between visual rating and leaf chlorophyll concentration. *Plant Breed.* 119, 365–367. doi: 10.1046/j.1439-0523.2000.00506.x
- Xu, X., Pan, S. K., Cheng, S. F., Zhang, B., Mu, D. S., Ni, P. X., et al. (2011). Genome sequence and analysis of the tuber crop potato. *Nature* 475, 189–194. doi: 10.1038/nature10158
- Yamori, W., Hikosaka, K., and Way, D. A. (2014). Temperature response of photosynthesis in C₃, C₄, and CAM plants: temperature acclimation and temperature adaptation. *Photosynth. Res.* 119, 101–117. doi: 10.1007/s11120-013-9874-6
- Yan, J. Q., He, C. X., Wang, J., Mao, Z. H., Holaday, S. A., Allen, R. D., et al. (2004). Overexpression of the *Arabidopsis* 14-3-3 protein GF14 lambda in cotton leads to a “Stay-Green” phenotype and improves stress tolerance under moderate drought conditions. *Plant Cell Physiol.* 45, 1007–1014. doi: 10.1093/pcp/pch115
- Yang, J., Zhang, J., Huang, Z., Zhu, Q., and Long, W. (2000). Remobilization of carbon reserves is improved by controlled soil-drying during grain filling of wheat. *Crop Sci.* 40, 1021–1023. doi: 10.2135/cropsci2000.4061645x

- Yang, J., Zhang, J., Zhiqing, W., Qingsen, Z., and Lijun, L. (2001). Water deficit-induced senescence and its relationship to the remobilization of pre-stored carbon in wheat during grain filling. *Agron. J.* 93, 196–206. doi: 10.2134/agronj2001.931196x
- Yu, D., Ranathunge, K., Huang, H., Pei, Z., Franke, R., Schreiber, L., et al. (2008). Wax crystal-sparse Leaf1 encodes a beta-ketoacyl CoA synthase involved in biosynthesis of cuticular waxes on rice leaf. *Planta* 228, 675–685. doi: 10.1007/s00425-008-0770-9
- Zhu, X. Y., and Xiong, L. Z. (2013). Putative megaenzyme DWA1 plays essential roles in drought resistance by regulating stress-induced wax deposition in rice. *Proc. Natl. Acad. Sci. U.S.A.* 110, 17790–17795. doi: 10.1073/pnas.1316412110

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ghatak, Chaturvedi, Bachmann, Valledor, Ramšak, Bazargani, Bajaj, Jegadeesan, Li, Sun, Gruden, Varshney and Weckwerth. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cabo Verde's Poaceae Flora: A Reservoir of Crop Wild Relatives Diversity for Crop Improvement

Vanézia Rocha^{1,2†}, Maria C. Duarte^{2†}, Sílvia Catarino^{1,3}, Ivani Duarte⁴ and Maria M. Romeiras^{1,2*}

¹ Linking Landscape, Environment, Agriculture and Food (LEAF), Instituto Superior de Agronomia (ISA), Universidade de Lisboa, Lisbon, Portugal, ² Centre for Ecology, Evolution and Environmental Changes (cE3c), Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, ³ Forest Research Centre (CEF), Instituto Superior de Agronomia (ISA), Universidade de Lisboa, Lisbon, Portugal, ⁴ Parque Natural do Monte Gordo, Ministério da Agricultura e Ambiente, Praia, Cabo Verde

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Arindam Ghatak,
University of Vienna, Austria
Mehanthan Muthamilarasan,
University of Hyderabad, India
Puneet Paul,
University of Nebraska–Lincoln,
United States

*Correspondence:

Maria M. Romeiras
mmromeiras@isa.ulisboa.pt

[†] These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 17 November 2020

Accepted: 08 January 2021

Published: 01 February 2021

Citation:

Rocha V, Duarte MC, Catarino S,
Duarte I and Romeiras MM (2021)
Cabo Verde's Poaceae Flora: A
Reservoir of Crop Wild Relatives
Diversity for Crop Improvement.
Front. Plant Sci. 12:630217.
doi: 10.3389/fpls.2021.630217

Africa is home to important centers of origin and diversity of crop wild relatives (CWR), including many species adapted to adverse agroecological conditions, namely drought and poor soils. Plant genetic resources from Cabo Verde Islands have been poorly explored for their potential to supplement the genetic pool of cultivated species. In this paper we identify Cabo Verde's CWR from the Poaceae family and provide a checklist of priority CWR *taxa*, highlighting those of particular conservation concern and the areas which should be the focus of the most intensive conservation efforts in these islands. Our results revealed that Cabo Verde archipelago is an important center of CWR diversity of West African crop millets, namely fonio (e.g., white fonio, *Digitaria exilis*, and black fonio, *Digitaria iburua*) and other African millets [e.g., pearl millet (*Cenchrus americanus* = *Pennisetum glaucum*), teff millet (*Eragrostis tef*), finger millet (*Eleusine coracana*), barnyard millet (*Echinochloa colona*), proso millet (*Panicum miliaceum*), and foxtail millet (*Setaria italica*)], which represent a diverse group of cereal crops, and important components in agriculture and food security of this country. Also, hotspot areas of diversity for *in situ* conservation were identified in Cabo Verde, as well as several populations occurring under extreme habitats conditions that are well adapted to drylands and poor soils. The evaluation of their potential for new ecologically important adaptive characteristics associated with tolerance to abiotic stresses is discussed. The survey of international Germplasm Banks revealed that very few accessions from Cabo Verde are conserved, contributing to the loss of genetic diversity of plant genetic resources in this archipelago. Particularly, the diversity of millets and the associated indigenous knowledge are critical for the food security and cultural identity of many poor farmers in Cabo Verde.

Keywords: conservation strategies, crop wild relative (CWR), millets, plant genetic resources, grass flora, oceanic island, prioritization

INTRODUCTION

Crop Wild Relatives (CWR) are those species growing in natural habitats that are genetically related to food, fodder and forage crops, medicinal plants, condiments, ornamental and forestry species used by humankind (Maxted et al., 2006). CWR are valuable sources of adaptive traits, contributing to biotic and abiotic stress tolerance, and thereby allow crop improvements with a wide range of agronomical and nutritional benefits (Castañeda-Álvarez et al., 2016). As actual or potential gene donors, CWR have a wide range of important genetic traits due to their widespread adaptation to different habitats and to continuous *in situ* evolution and because they have not undergone domestication processes (Vincent et al., 2013; Zhang et al., 2017).

Considering an increasing population, the conservation of CWR diversity to ensure food security is of paramount importance and has been widely recognized by several world organizations, namely the Food and Agriculture Organization (Vincent et al., 2013). The Global Crop Diversity Trust has proposed projects like the “Global initiative to collect, conserve, and use crop wild relatives” (Dempewolf et al., 2014), and the 2030 agenda of Sustainable Development Goals (SDGs), implemented in 2015 by the United Nations, advocates the promotion of sustainable agriculture (SDGs2) by reinforcing resilience and adaptation to climate change (SDGs13) and preventing desertification and protection of biodiversity (domesticated and wild species) (SDGs15) (Sonesson et al., 2016). Like other Plant Genetic Resources, CWR are mainly threatened by the impacts of anthropogenic activities, such as habitat destruction, pollution and urbanization, and even competition by invasive species, but climate changes seem to be a particularly serious problem for this group (Ford-Lloyd et al., 2011; van Treuren et al., 2017; Teso et al., 2018; Allen et al., 2019). Some CWR may be passively conserved *in situ* due to strategies targeting other species, namely in protected areas, but specific conservation linked to CWR remains necessary.

The need for CWR conservation has been widely discussed (Heywood et al., 2007; Maxted et al., 2013) and the CWR inventory seems to be a crucial step to identify the conservation requirements (Zhang et al., 2017; Teso et al., 2018). Approaches to protect CWR diversity can involve an individual, national, regional or global level planning, with prioritization criteria that should ensure a successful conservation strategy with limited resources for implementation (Allen et al., 2019). The prioritization criteria can be adjusted according to the conservation strategy to be adopted, however, the main ones are: (i) economic value of the related crop, mainly its importance for human food and livestock supply; (ii) genetic potential as gene donor, priority being given to CWR that are more closely related to the crop; (iii) occurrence status of the CWR, as native, introduced or invasive on the geographical area in question; and iv) threat status (Maxted et al., 2013). Moreover, the Gene Pool (GP) (Harlan and de Wet, 1971) and Taxonomic Group (TG) (Maxted et al., 2006) concepts are used to determine the crossability between a crop and CWR. They are based on genetic and taxonomical relatedness, respectively GP and TG,

and determine how closely a CWR relates to a crop and how easily they can cross (Maxted et al., 2006; Vincent et al., 2013).

The Poaceae family, with the third highest global priority among crop wild relatives, accounts for ca. 150 priority CWR distributed over 18 genera [*Aegilops* (= *Amblyopyrum*), *Agropyron*, *Avena*, *Cenchrus* (= *Pennisetum*), *Digitaria*, *Echinochloa*, *Eleusine*, *Elymus*, *Hordeum*, *Oryza*, *Panicum*, *Saccharum*, *Secale*, *Setaria*, *Sorghum*, *Tripsacum*, *Triticum*, and *Zea*] which, all except *Aegilops* and *Tripsacum*, belong to the global priority list of 92 CWR genera (Vincent et al., 2013). Among these genera are found most of the species that are wild relatives of the main consumption cereal crops (e.g., rice, wheat, maize, and oat), and other Poaceae food crops, such as sugarcane and sorghum, that substantially contribute to the human dietary energy (Vincent et al., 2013; Allen et al., 2019). Furthermore, in the semi-arid tropical regions of Africa and Asia, millets (small grain crops) are important sources of energy and protein for millions of persons living in developing countries (Amadou et al., 2013). The millets' group comprises many different species, with pearl millet (*Cenchrus americanus* = *Pennisetum glaucum*), foxtail millet (*Setaria italica*), proso millet (*Panicum miliaceum*), and finger millet (*Eleusine coracana*) being the most important ones (Dwivedi et al., 2012; Amadou et al., 2013; Tadele, 2016). The maintenance of African millet diversity depends on agricultural, food and livelihood dynamics at the farmer level, since every community holds local cultivars to address their agroecological conditions, farming practices, and food needs (IPGRI, 2002). Small millets of the Poaceae family have been commonly mentioned as ‘smart foods’ or ‘nutri-cereals’ because they are more efficient in water and nitrogen use than major cereals like rice or maize (Muthamilarasan and Prasad, 2021). They grow in a diverse range of environmental conditions, as they are more tolerant to diseases, pests, and abiotic stresses (Vetriventhan et al., 2020).

To secure the long-term conservation, worldwide genebanks hold accessions of cultivated and wild germplasm of small millets, but only of the most important species, as is the case of finger millet, foxtail millet, and proso millet (Muthamilarasan and Prasad, 2021). Recently, Varshney et al. (2017) report the whole-genome sequence of pearl millet, providing an important resource to improve agronomic traits in arid environments and accelerate millets crop improvement.

Millets are well adapted to adverse climatic conditions (limited rainfall) and are mainly cultivated in marginal agricultural areas (Tadele, 2016), playing a major role in rural agriculture of West African countries, and particularly in the tropical dry islands of Cabo Verde (Teixeira and Barbosa, 1958). Although its agriculture is limited by a set of natural constraints (e.g., persistent drought periods, scarcity of quality soil, small territory available as farmland), this archipelago has shown considerable progresses toward overall development in the agriculture sector during the last two decades (Varela et al., 2020). Nevertheless, there is only official information on cultivation for two Poaceae species, *Zea mays* (maize) and *Saccharum officinarum* (sugarcane) (Monteiro et al., 2020). Maize is cultivated as a rainfed crop and used for human food and fodder; sugarcane is the most important irrigated crop and occupies the largest

harvest area, its main purpose being the production of a by-product, the highly alcoholic drink “grog,” very much appreciated and exported as a national product (Monteiro et al., 2020).

Therefore, and despite the great importance of the Poaceae family as the dominant element of the native flora of Cabo Verde, the potential of plant genetic resources of these islands to supplement the genetic pool of cultivated species has been poorly explored. Considering the huge wealth of natural grass populations occurring in these islands under extreme environmental conditions, from dry lowlands to less dry altitude zones or even to subalpine regions (Neto et al., 2020), assessing the diversity of wild species with potential adaptive characteristics for abiotic stress tolerance would generate new data on varieties of Cabo Verde adapted to drought conditions (Essoh et al., 2020). This could contribute to valorise plant genetic resources and, in the future, to generate new eco-products with national economic impacts at the industry and economy levels.

Although this is one of the most important plant families in Cabo Verde and despite the importance of several crop species for humans and livestock, studies are scarce, particularly concerning the diversity of grasses in this archipelago. Moreover, there are threats, particularly to the endemic flora (Romeiras et al., 2016), requiring *in situ* (in natural habitats) and *ex situ* (in gene banks) conservation of the unique plant genetic resources that can assist the improvement of these crops and, consequently, ensure food security (Monteiro et al., 2020).

In this paper we intend to identify Cabo Verde's CWR from the Poaceae family and to provide a checklist of priority CWR *taxa*, highlighting those of particular conservation concern. Based on a gap analysis, hotspot areas of Poaceae CWR diversity will be identified for this archipelago, in order to provide new data to propose future *in situ* conservation actions. Also, the total number of *ex situ* accessions will be identified for the African continent and, specifically, for Cabo Verde, in order to support future management of seed collection and conservation of local plant genetic resources. Finally, for the priority CWR Poaceae species identified, a comparative global analysis will be performed, and the results will be discussed in the context of West African food security.

MATERIALS AND METHODS

Study Area

Cabo Verde is a North Atlantic Ocean archipelago and corresponds to the southernmost islands of Macaronesia. It is located at latitudes 14°45′ – 17°10′ N and longitudes 22°40′ – 25°20′ W, about 1,350 km south-west of the Canary Islands and ca. 560 km from Senegal's coast. The archipelago (see **Figure 1**) has a total area of ca. 4,033 km² and includes ten major islands distributed in three groups: Northern Islands [Santo Antão, São Vicente, Santa Luzia (the only uninhabited island), and São Nicolau]; Eastern Islands (Sal, Boavista and Maio); and the Southern Islands (Santiago, Fogo and Brava) (Duarte and Romeiras, 2009). Currently, the population is estimated at 556 thousand inhabitants but is expected to reach 679 thousand in 2050 (United Nations [UN], 2019).

Cabo Verde has a dry tropical climate with two well-marked seasons conditioning the distribution of its flora and vegetation (Neto et al., 2020). The long dry season of 8–10 months varies between November and July, and the sparse and irregular humid rainy season of 1–3 months, usually occurs from August to October (Monteiro et al., 2020). Mean annual temperature is usually around 25°C with low thermal amplitude, due to maritime influence; lowest monthly average temperatures normally occur in January or February (ca. 18°C), and the highest in September (ca. 27°C), however, temperatures as high as 35–40°C can occur in internal regions of the arid Eastern Islands (Duarte and Romeiras, 2009). The average annual relative air humidity ranges from 75% to more than 80% (Monteiro et al., 2020). Particularly, the topography of the islands contributes to significant spatial variations of flora and vegetation, with altitude and exposure to northeast trade winds leading to contrasting weather conditions (Duarte et al., 2008). The northern and southern islands are characterized by high mountains, offering a wide range of habitats over relatively short distances, whereas the eastern islands are lower and drier (Romeiras et al., 2015). Particularly, the lowland flora of Cabo Verde is markedly of Afrotropical origin and dominated by grass species, whereas the endemic mountain flora that mainly occurs on the north/northeast-facing slopes shows affinities with Madeiran and Canarian flora (Rivas-Martínez et al., 2017; Freitas et al., 2019).

Poaceae Crop Wild Relatives' Inventory and Data Collection

The Poaceae checklist from Cabo Verde was mainly based on data collected in previous fieldwork and inventories performed over the last two decades by M.C. Duarte, and on data from the literature (e.g., Lobin, 1986; Arechavaleta et al., 2005). Further data was obtained from specimens housed in worldwide herbaria (e.g., LISU, COI, K, and P), from Cabo Verde collections (e.g., in Parque Natural do Monte Gordo, São Nicolau Island), and information available in the Global Biodiversity Information Facility website (GBIF.org, 2020).

After completing the checklist of Cabo Verde CWR species, specific information was gathered on: (i) their status in Cabo Verde Islands - native or introduced; (ii) their distribution in the archipelago - number of islands where the species occur; (iii) their area of occupancy (AOO) calculated with a 1 km² grid using occurrences of each species with valid coordinates, with QGIS v.3.10.5 software (QGIS Development Team, 2020); (iv) their world distribution; (v) their confirmed and potential traits and respective gene pool or taxon group, based on the Crop Wild Relative Project database (CWR, 2019); and (vi) their ethnobotany, according to information from GRIN - Germplasm Resources Information Network (USDA, 2020), Plant Resources of Tropical Africa (PROTA4U, 2020) and Cabo Verde herbarium data. Taxonomical nomenclature was checked according to Plants of the World Online - The Plant List (2013) and POWO (2019).

Crop species and their scientific names were gathered from the available literature (Dwivedi et al., 2012; Amadou et al., 2013; Vincent et al., 2013; Tadele, 2016). The native distribution

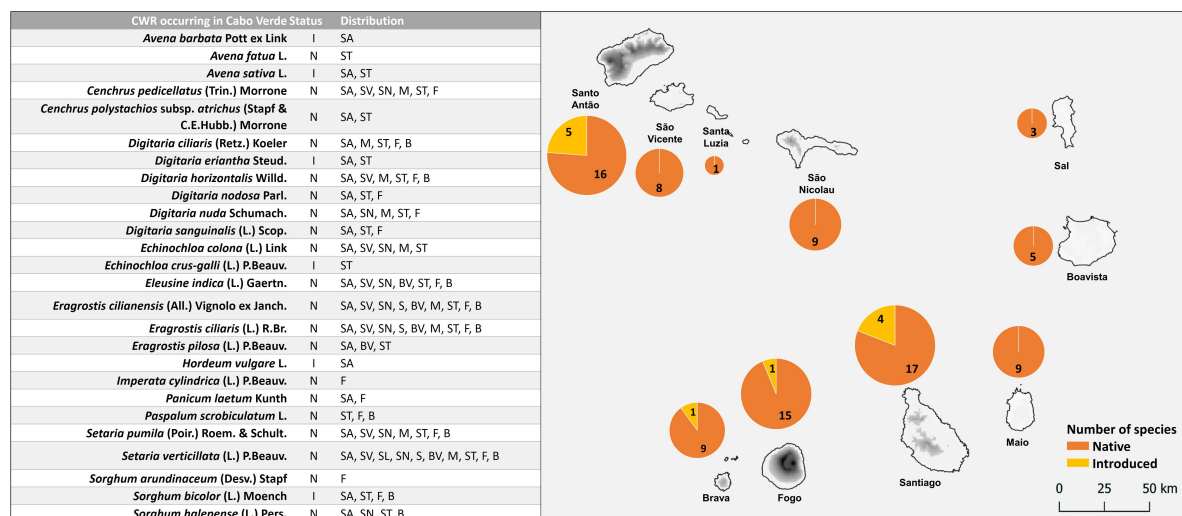


FIGURE 1 | Crop wild relatives (CWR) Poaceae species occurring in Cabo Verde: status and distribution in the archipelago. The area of the circles is proportional to the number of species. Status: I, introduced; N, native. Island abbreviations: SA, Santo Antão; SV, São Vicente; SL, Santa Luzia; SN, São Nicolau; S, Sal; BV, Boavista; M, Maio; ST, Santiago; F, Fogo; B, Brava.

of each crop species was obtained from GRIN (USDA, 2020), preferentially, and from POWO (2019). Additionally, for each crop species two measures of food supply – protein (g/100 g) and fat (g/100 g) – were retrieved from available literature (Kajuna, 2001; Belton and Taylor, 2004; Amadou et al., 2013; Chandi and Annor, 2016; Taylor, 2016; Williams et al., 2016; Wrigley, 2016) and the ICRISAT website (ICRISAT, 2020); also, two measures of global agricultural production – harvested area (ha) and production quantity (tons), using data from three recent years (2016–2018) (Food and Agricultural Organization of the United Nations [FAOSTAT], 2020) – were included. When the information on the agricultural production was aggregated for the listed crop, which is usually the case of millets, the values were disaggregated through division of the total value by all the identified crops.

Importance and Priority Scores

For the Cabo Verde Poaceae wild relatives, the genetic potential as gene donor of each CWR species was identified, based on the gene pool concept (sublevel: GP1a, GP1b, GP2, and GP3) or, when this information was unavailable, on the taxon group concept (sublevel: TG1a, TG1b, TG2, TG3, TG4, and TG5). According to Maxted et al. (2006) and Vincent et al. (2013), the highest priority CWR is the one that can most easily cross with the crop, namely GP1b, GP2, when using gene pool concept or TG1b, TG2, and TG3, in the case of the taxon group concept. The genetic potential as gene donor is one of the most used parameters to establish conservation priorities.

The importance of each associated crop was estimated by applying the “Importance Score” (IS), adapted from Castañeda-Álvarez et al. (2016). The Importance Score was produced: initially, by dividing the food supply (protein and fat) and the agricultural production (harvested area and production quantity) by the maximum existing value across all crops;

then, the food supply and agricultural production metrics were averaged separately; finally, the Importance Score was produced by averaging the mean food supply and mean agricultural production values (Castañeda-Álvarez et al., 2016).

In this study, a “Priority Score” (PS) for each CWR was calculated based on nine criteria (see **Table 1**), two directly related with the associated crops (I and II) and seven with the crop wild relatives themselves (III to IX). The prioritization criteria were adapted from Heywood et al. (2007) and Maxted et al. (2013). For each class, within each criterion, a value of 1–3 was assigned concerning conservation importance (1 – low, 2 – medium, 3 – high) as follows: (I) crop Importance Score – 1 ($IS \leq 0.2$), 2 ($0.2 < IS \leq 0.4$), 3 ($IS > 0.4$); (II) native distribution of the crop – 1 (native out Africa), 2 (native to Africa, excluding West Africa), 3 (native to West Africa and other regions); (III) genetic potential as gene donor of the CWR – 1 (GP3 or TG3), 2 (GP2 or TG2), 3 (GP1 or TG1); (IV) number of associated crops – 1, 2, or 3 (respectively, with one, two or three crops); (V) status in Cabo Verde – 1 (introduced, native from regions out of Africa), 2 (introduced, native in Africa but not in Cabo Verde), 3 (native in Cabo Verde); (VI) distribution in Cabo Verde – 1 (≥ 6 islands), 2 (4–5 islands), 3 (1–3 islands); (VII) area of occupancy, as a proxy to the threat status in Cabo Verde – 1 ($> 40 \text{ km}^2$), 2 ($21 - 40 \text{ km}^2$), 3 ($\leq 20 \text{ km}^2$); (VIII) world distribution – 1 (native out Africa), 2 (native to Africa, excluding West Africa), 3 (native to West Africa, and other regions); (IX) ethnobotanical uses – 1 (≤ 2 uses), 2 (3–4 uses), 3 (> 4 uses). For each CWR species, PS results from the sum of all the values assigned to it and priority classes were established as follows: low – up to 18; medium – up to 21; high – higher than 21.

For the CWR species related to more than one crop (as, for instance, *Echinochloa colona* to barnyard millet and Indian barnyard millet, see **Table 2**), the Priority Score (see **Table 3**) was calculated considering, for each criterion, the highest assigned

TABLE 1 | Summary of the nine criteria and the scores assigned to each CWR and its associated crop.

	No.	Criteria	Score 1 - Low	Score 2 - Medium	Score 3 - High
Associated crop	I	Crop Importance Score (IS)	IS ≤ 0.2	0.2 < IS ≤ 0.4	IS > 0.4
	II	World native distribution of the crop	Native out Africa	Native to Africa, excluding West Africa	Native to West Africa and other regions
CWR	III	CWR genetic potential as gene donor	GP3 or TG3	GP2 or TG2	GP1 or TG1)
	IV	Number of associated crops	1	2	3
	IV	CWR status in Cabo Verde	Introduced, native from regions out of Africa	Introduced, native in Africa but not in Cabo Verde	Native in Cabo Verde
	VI	Distribution in Cabo Verde (number of islands)	≥6 islands	4–5 islands	1–3 islands
	VII	Area of occupancy (AOO)	>40 km ²	21–40 km ²	≤20 km ²
	VIII	World native distribution	Native out Africa	Native to Africa, excluding West Africa	Native to West Africa and other regions
	IX	Ethnobotanical uses	≤2 uses	3–4 uses	>4 uses

score (e.g., if a CWR belongs to GP1 of one crop and to GP3 of another, a score of 3 was given). For comparative analyzes of crops, the mean PS was gathered adding the individual values of all the CWR associated to a given crop and dividing the total by the corresponding CWR number (Castañeda-Álvarez et al., 2016).

Diversity Hotspots and *in situ* Conservation Gap Analysis

The distribution of Poaceae CWR in Cabo Verde was estimated with occurrence data collected from previous fieldwork, studied specimens housed in worldwide herbaria and information available at the GBIF website (GBIF.org, 2020). Specimens without geographical coordinates were georeferenced following the Guide to Best Practices for Georeferencing (Chapman and Wiecek, 2006) and using Google Earth Pro 7.3.2.5491 (Sere, 2018). Duplicate records, i.e., with the same collector and the same number of collection, were excluded. The final dataset with 675 occurrences with geographical coordinates was used in the analysis.

The altitudinal distribution of each CWR species was estimated based on the interception of the occurrence records with the map of altitude provided by CGIAR-CSI Consortium for Spatial Information (CGIAR-CSI Consortium for Spatial Information, 2020) at a resolution of 90 m. To summarize these data, we built a boxplot graph, using R version 3.6.0. (R Development Core Team, 2020).

Based on georeferenced occurrence data, we also constructed a species richness map using QGIS v.3.10.5 (QGIS Development Team, 2020). This map presents the number of species occurring in each cell of 4 km², allowing to identify the areas of greatest diversity of CWR species in Cabo Verde. Species occurrences and diversity hotspots were then overlaid with the national network of protected areas of Cabo Verde, downloaded from Infra-estrutura de Dados Espaciais de Cabo Verde (IDE-CV, 2020). This analysis aimed to assess the coverage and efficiency of the network of protected areas to preserve CWR species, identifying the main conservation gaps.

Ex situ Conservation Analysis

The status of *ex situ* conservation of Cabo Verde's Poaceae CWR in worldwide genebanks was assessed through the Genesys Database (Genesys, 2020). Comprehensive data for West Africa and specifically for Cabo Verde were identified in order to support future management of seed collection and conservation of local plant genetic resources.

RESULTS

Diversity of Cabo Verde Poaceae CWR

The inventory of the Poaceae family revealed that ca. 123 native and introduced *taxa* occur in Cabo Verde Islands. Twenty-six species are CWR, including five native species (*Eleusine indica*, *Eragrostis cilianensis*, *Eragrostis ciliaris*, *Setaria pumila*, and *Setaria verticillata*) that are widespread in the archipelago, occurring in more than 7 islands, and seven species (*Avena barbata*, *Avena sativa*, *Digitaria eriantha*, *Hordeum vulgare*, *Imperata cylindrica*, *Panicum laetum*, and *Sorghum arundinaceum*) that occur only in a single island (Figure 1). Santiago has the highest diversity of species (21), including 17 native and four introduced ones (Figure 1). Santo Antão, also with 21 CWR, is the island hosting more introduced species (*A. barbata*, *A. sativa*, *D. eriantha*, *H. vulgare* and *Sorghum bicolor*). The lowest number of CWR was found in Sal (3 species), and Santa Luzia (1 species – *S. verticillata*).

Among the CWR species, six are the wild forms of the identified crop (i.e., *A. sativa*, *E. colona*, *Echinochloa crus-galli*, *H. vulgare*, *Paspalum scrobiculatum*, and *S. bicolor*) and three (*E. colona*, *E. crus-galli*, and *S. bicolor*) are associated with more than one crop. Of the CWR occurring in Cabo Verde, twenty are native to the islands and the remaining six *taxa* are introduced; four of the latter have a native distribution range in Africa (Table 2, more details in Supplementary Table 1). Three crops (fonio, pearl millet, and sorghum) have native distributions exclusively in the African continent, and

TABLE 2 | List of **(A)** Poaceae crops considered in this study, their Importance Score and native distributions; and **(B)** the associated crop wild relatives occurring in Cabo Verde Islands and information on their gene pool, native status, and uses.

(A) Crops				(B) Associated crop wild relatives			
Crop name	Scientific name	Importance Score	Native distribution	CWR occurring in Cabo Verde	Gene pool	Status in Cabo Verde	Uses
Barley	<i>Hordeum vulgare</i> L.	0.504	Africa (N); Europe; Asia	<i>Hordeum vulgare</i> *	GP1	I	H, F, Fo, E, M, Me
Barnyard millet	<i>Echinochloa colona</i> (L.) Link.	0.304	Africa; Middle East; South Asia. Europe (E)	<i>Echinochloa colona</i> *	GP1	N	H, F ^{CV} , Fo, E
	<i>Echinochloa crus-galli</i> (L.) P.Beauv.			<i>Echinochloa crus-galli</i> *	GP1	I	H, F, E, Me
Finger millet	<i>Eleusine coracana</i> subsp. <i>coracana</i> (L.) Gaertn.	0.189	Africa; Asia (Temperate)	<i>Eleusine indica</i>	GP1	N	H, F ^{CV} , Fo, M, Me
Fonio	<i>Digitaria exilis</i> (Kippist) Stapf (white fonio) & <i>Digitaria iburua</i> Stapf (black fonio)	0.304	Africa (W)	<i>Digitaria ciliaris</i>	GP3	N	H, F ^{CV} , E
				<i>Digitaria eriantha</i>	GP3	I	F, Fo, E, M, Me
				<i>Digitaria horizontalis</i>	GP3	N	H, F
				<i>Digitaria nodosa</i>	GP3	N	F ^{CV}
				<i>Digitaria nuda</i>	GP3	N	H, F
				<i>Digitaria sanguinalis</i>	GP3	N	F, Fo, E
Foxtail millet	<i>Setaria italica</i> (L.) P.Beauv.	0.383	Africa (N); Europe; Asia	<i>Setaria verticillata</i>	GP2	N	H, F ^{CV} , Fo, E, M, Me
				<i>Setaria pumila</i>	GP3	N	H, F ^{CV} , E, M, Me
Indian barnyard millet	<i>Echinochloa frumentacea</i> Link.	0.203	Asia (S)	<i>Echinochloa colona</i>	GP1	N	H, F ^{CV} , Fo, E
				<i>Echinochloa crus-galli</i>	GP3	I	H, F, E, Me
Japanese barnyard millet	<i>Echinochloa esculenta</i> (A.Braun) H.Scholz	0.203	Asia (E)	<i>Echinochloa crus-galli</i>	GP1	I	H, F, E, Me
Kodo millet	<i>Paspalum scrobiculatum</i> L.	0.210	Africa; Pacific; Asia; Australia	<i>Paspalum scrobiculatum</i> *	GP1	N	H, F, Fo, E, M, Me
Oat	<i>Avena sativa</i> L.	0.537	Middle East	<i>Avena sativa</i> *	GP1	I	H, F, Fo, Me
				<i>Avena fatua</i>	GP1	N	F, Fo, Me
				<i>Avena barbata</i>	GP3	I	F
Pearl millet	<i>Cenchrus americanus</i> (L.) Morrone	0.449	Africa (W, C, S)	<i>Cenchrus pedicellatus</i>	GP3	N	F, Fo, M, Me
				<i>Cenchrus polystachios</i> subsp. <i>atrichus</i>	GP3	N	H, F, E, M, Me
Proso millet	<i>Panicum miliaceum</i> L.	0.272	Asia (Temperate)	<i>Panicum laetum</i>	GP3	N	H, F, Me
Sorghum	<i>Sorghum bicolor</i> (L.) Moench	0.469	Africa	<i>Sorghum bicolor</i> *	GP1	I	H, F, Fo, E, M, Me
				<i>Sorghum arundinaceum</i>	GP1	N	H, F ^{CV} , M, Me, O
				<i>Sorghum halepense</i>	GP2	N	H, F, Me
Sugarcane	<i>Saccharum officinarum</i> L.	0.502	Oceania (New Guinea)	<i>Imperata cylindrica</i>	GP3	N	H, F, E, M, Me, O
				<i>Sorghum bicolor</i>	GP3	I	H, F, Fo, E, M, Me
Teff (millet)	<i>Eragrostis tef</i> (Zuccagni) Trotter	0.254	Africa (E); Middle East	<i>Eragrostis pilosa</i>	GP1	N	H, F, Me
				<i>Eragrostis cilianensis</i>	TG2 ^a	N	H, F ^{CV} , Fo, E, M
				<i>Eragrostis ciliaris</i>	TG2 ^a	N	H, F ^{CV} , M, Me

CWR taxon: *The wild form of the crop. Status: I, introduced; N, native. CWR uses: H, human consumption; F, forage; Fo, fodder; E, environmental; M, materials; Me, medicinal; O, other; ^{CV}, with information reported for Cabo Verde. (a) according to Ingram and Doyle (2007).

TABLE 3 | Poaceae CWR occurring in Cabo Verde, their native status, distribution, area of occupancy, gene pool, scoring of the prioritization criteria and conservation priority levels.

CWR taxon	Native status	Number of islands	Area of occupancy (km ²)	Gene pool (higher level)	Criteria and respective scores ^a									Priority Score	Priority category
					Associated crop		CWR								
							I	II	III	IV	V	VI	VII		
<i>Avena barbata</i>	I	1	1	GP3	3	1	1	1	1	3	3	2	1	16	Low
<i>Avena fatua</i>	N	1	< 0.5	GP1	3	1	3	1	3	3	3	2	2	21	Medium
<i>Avena sativa</i>	I	2	3	GP1	3	1	3	1	1	3	3	1	2	18	Low
<i>Cenchrus pedicellatus</i>	N	6	23	GP3	3	3	1	1	3	1	2	3	2	19	Medium
<i>Cenchrus polystachios</i> subsp. <i>atrichus</i>	N	2	3	GP3	3	3	1	1	3	3	3	3	3	23	High
<i>Digitaria ciliaris</i>	N	5	50	GP3	2	3	1	1	3	2	1	3	2	18	Low
<i>Digitaria eriantha</i>	I	2	1	GP3	2	3	1	1	2	3	3	2	3	20	Medium
<i>Digitaria horizontalis</i>	N	6	21	GP3	2	3	1	1	3	1	2	1	1	15	Low
<i>Digitaria nodosa</i>	N	3	11	GP3	2	3	1	1	3	3	3	3	1	20	Medium
<i>Digitaria nuda</i>	N	5	56	GP3	2	3	1	1	3	2	1	3	1	17	Low
<i>Digitaria sanguinalis</i>	N	3	2	GP3	2	3	1	1	3	3	3	2	2	20	Medium
<i>Echinochloa colona</i>	N	5	35	GP1	2	1	3	2	3	2	2	3	2	20	Medium
<i>Echinochloa crus-galli</i>	I	1	< 0.5	GP1	2	1	3	3	1	3	3	1	2	19	Medium
<i>Eleusine indica</i>	N	7	54	GP1	1	3	3	1	3	1	1	3	3	19	Medium
<i>Eragrostis cilianensis</i>	N	9	45	TG2	2	2	2	1	3	1	1	3	3	18	Low
<i>Eragrostis ciliaris</i>	N	9	38	TG2	2	2	2	1	3	1	2	3	2	18	Low
<i>Eragrostis pilosa</i>	N	3	< 0.5	GP1	2	2	3	1	3	3	3	3	2	22	High
<i>Hordeum vulgare</i>	I	1	2	GP1	3	2	3	1	1	3	3	2	3	21	Medium
<i>Imperata cylindrica</i>	N	1	3	GP3	3	1	1	1	3	3	3	3	3	21	Medium
<i>Panicum laetum</i>	N	2	1	GP3	2	1	1	1	3	3	3	3	2	19	Medium
<i>Paspalum scrobiculatum</i>	N	3	14	GP1	2	3	3	1	3	3	3	3	3	24	High
<i>Setaria pumila</i>	N	7	40	GP3	2	2	1	1	3	1	2	3	3	18	Low
<i>Setaria verticillata</i>	N	10	75	GP2	2	2	2	1	3	1	1	2	3	17	Low
<i>Sorghum arundinaceum</i>	N	1	1	GP1	3	3	3	1	3	3	3	3	3	25	High
<i>Sorghum bicolor</i>	I	4	7	GP1	3	3	3	2	2	2	3	3	3	24	High
<i>Sorghum halepense</i>	N	4	20	GP2	3	3	2	1	3	2	3	2	2	21	Medium

^aCriteria: (I) crop importance; (II) world native distribution of the crop; (III) CWR genetic potential as gene donor; (IV) number of associated crops; (V) CWR status in Cabo Verde; (VI) distribution in Cabo Verde (number of islands); (VII) area of occupancy; (VIII) world native distribution; (IX) ethnobotanical uses.

six (barley, barnyard millet, finger millet, foxtail millet, kodo millet, and teff) have a native range that includes Africa; the other five crops (Indian barnyard millet, Japanese barnyard millet, oat, proso millet, and sugarcane) are non-native to Africa (Table 2).

Based on 675 occurrence records (Supplementary Table 2), we analyzed the altitudinal distribution of the studied CWR species; it ranges from sea level (*D. ciliaris*, *E. cilianensis*, and *S. pumila*) to 1,780 m (*E. ciliaris*) (Figure 2). Most species have a large altitudinal distribution, occurring from low altitude to more than 1,000 m. The species *Digitaria nodosa*, *E. colona*, and *E. cilianensis* were identified mainly in lower areas, with median altitudes of 320, 201, and 313 m, respectively. The species *H. vulgare* and *I. cylindrica* stand out for their ability to grow at high altitude, with medians of 1,306 and 1,560 m, respectively.

Crop Importance and CWR Priority Scores

Among the Poaceae crops considered in this study, five can be considered of high importance according to the Importance Score (Table 2): pearl millet, sorghum, sugarcane, barley, and oat. Eight crops are of medium importance: Indian barnyard millet, Japanese barnyard millet, kodo millet, teff (millet), proso millet, barnyard millet, fonio, and foxtail millet. Only the finger millet was classified as of low crop importance.

The “Priority Scores” (PS) for crop wild relatives, ranging from 15 to 25, with score 15 corresponding to low priority CWR (*Digitaria horizontalis*), and CWR with the highest priority (score 25) being *S. arundinaceum* (Table 3, Figure 3, and Supplementary Table 3). From the 26 identified CWR, five (19.2%) taxa were ranked as high priority, 12 (46.2%) as medium priority and nine (34.6%) as low priority (Figure 3). In terms of gene pool, 10 taxa were classified as GP1 (38.5%), four as GP2 and TG2 (15.4%), and 12 as GP3 (46.1%) (Table 3). Three species are wild relative to more than one crop: *E. crus-galli*, GP1 of barnyard millet and Japanese barnyard millet, and GP3 of Indian barnyard millet; *E. colona*, GP1 of barnyard millet and Indian barnyard millet; and *S. bicolor*, GP1 of sorghum and GP3 of sugarcane.

Except for Boavista, mountain islands (maximum altitude between 900 and 2,800 m) are those with high-priority taxa (Figure 4 and Supplementary Table 4) such as *S. arundinaceum*, *S. bicolor*, *P. scrobiculatum*, *Cenchrus polystachios* subsp. *atrichum* and *Eragrostis pilosa*. In Santa Luzia only one CWR occurs (*S. verticillata*); this may be related to the fact that this island has been subject of less botanical exploration over the last decades than the inhabited islands, as well as to fewer available habitats due to its limited area.

A comparative view of the importance of Poaceae crops and their associated CWR in Cabo Verde is represented in Figure 5 (more details in Supplementary Table 5). Among the nine crops with native distribution in the African continent, six occur in West Africa, with kodo millet, sorghum and pearl millet being those with more associated high-priority CWR

[*P. scrobiculatum* (kodo millet); *S. arundinaceum* and *S. bicolor* (sorghum); *C. polystachios* subsp. *atrichus* (pearl millet)]. Among priority species to collect and conserve, the CWR of sorghum and pearl millet should be highlighted (upper right part of Figure 5: PS > 21; IS > 0.4). Fonio is the crop with more associated CWR in Cabo Verde (6), with a native distribution exclusive to West Africa and all the associated CWR being used as forage (Table 2); however, the low Importance Score of the crop together with the fact that all the fonio CWR present in Cabo Verde belong to GP3, place this group as of low priority. Although this crop presents a medium Importance Score (0.304) it is an important African crop, which supports animal livestock and some human supply. Oat, sugarcane and barley, non-native to Cabo Verde, are among the most important crops; their associated CWR in Cabo Verde are mainly introduced species, with native distributions in Africa.

The CWR species themselves have important uses. We identified seven species with six different uses, six species with five uses, and eight species with four uses; nine species have three or less uses (Table 2). The most common use is as fodder (26 taxa), followed by human consumption (20 taxa) and medicinal applications (18 taxa).

Information regarding confirmed and potential traits are only available for barley, oat and sorghum crops, and mainly concerning abiotic, agronomic and biotic traits, such as drought tolerance, yield improvement and pathogen resistance, respectively. For the remaining crops, and especially millets, no information is available.

In situ Conservation: Hotspots and Conservation Gap Analysis

The main diversity hotspots are found in Santiago and Santo Antão (Figure 6). The maximum of eight species per cell of 2 km × 2 km was found in Santiago, between Pico da Antónia and Rui Vaz, including one high priority (*P. scrobiculatum*) and two medium priority species (*Cenchrus pedicellatus* and *E. indica*). Another important hotspot was found in the same island, near Ponta de Santa Cruz, with seven species, including one of high conservation priority (*P. scrobiculatum*) and four of medium priority (*C. pedicellatus*, *E. colona*, *E. indica*, and *S. halepense*). In Santo Antão, the highest diversity was found in Paúl (northeast of the island), with six species, two of them of medium conservation priority (*D. nodosa* and *E. colona*).

The most diverse areas in terms of species richness are generally outside the protected areas (Figure 6). Only 8.1% of the occurrence records, corresponding to 18 species (three species of high priority, eight of medium priority, and seven of low priority), are found within protected areas. Cova-Paúl-Ribeira da Torre Natural Park, in Santo Antão, hosts the highest number of CWR with seven species, however, only one of them is of high conservation priority and one is of medium priority (*S. bicolor* and *D. nodosa*, respectively). Six taxa, including two of high priority (*C. polystachios* subsp. *atrichus* and *P. scrobiculatum*) and two of medium priority (*D. sanguinalis* and *E. indica*), are found in Serra da Malagueta Natural Park. Five CWR species occur in Monte Verde Natural Park, three of them are medium

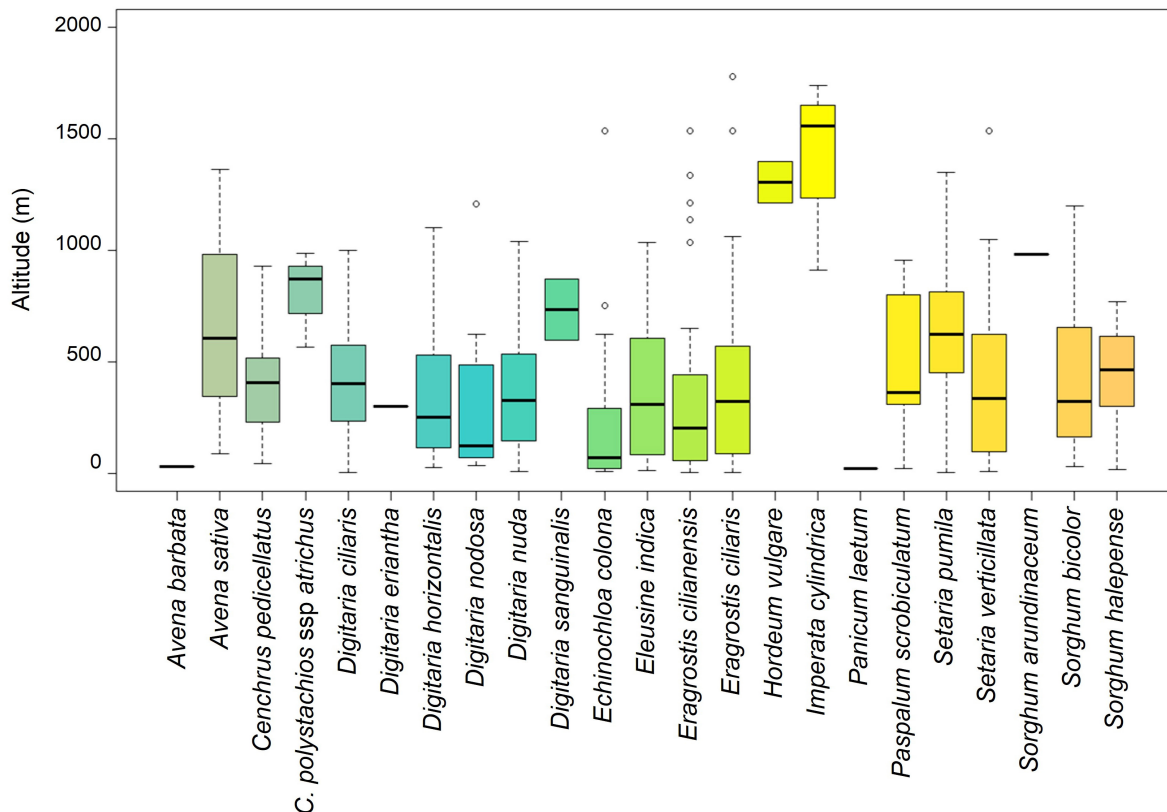


FIGURE 2 | Altitudinal distribution of CWR records in Cabo Verde Islands. The boxplots show the minimum (at the bottom of the chart, at the end of the vertical line), first quartile, Q1, (the bottom edge of the box), the median (line in the center of the box), third quartile, Q3 (the top edge of the box), the maximum (at the top end of the vertical line), and outliers (circles). Species represented with a single line have only one recorded occurrence. Species without records (*Avena fatua*, *Echinochloa crus-galli* and *Eragrostis pilosa*) were not included in this figure.

priority (*C. pedicellatus*, *E. colona*, and *E. indica*). Eight species – *A. barbata*, *A. fatua*, *D. eriantha*, *D. horizontalis*, *E. crus-galli*, *E. pilosa* (high priority), *P. laetum*, and *S. arundinaceum* (high priority) –, were not found in protected areas.

Ex situ Conservation

The analyses of the accessions present in global genebanks reveal that 25 species (96.1%) of CWR mentioned in the present work are currently conserved *ex situ* (Table 4). Considering the accessions collected worldwide, *H. vulgare* (252,545 accessions), a medium priority species, is the CWR with the highest number of accessions, followed by the high priority species *S. bicolor* (123,473 accessions), and the low priority species *Avena sativa* (53,617 accessions). Three taxa (11.5%) have 1,000 – 2,000 accessions, seven taxa (26.9%) have 100 – 1,000 accessions, and 12 taxa (46.2%) have 1 – 100 accessions. The only species without accession is *Digitaria nuda*.

Most high priority species are generally well represented in world genebanks: *S. arundinaceum* has a total of 430 accessions, but only 11 were collected in West Africa, namely in Mali; *S. bicolor* has more than 120,000 accessions; *P. scrobiculatum* has 1,114 accessions, 13 of them collected in West Africa; *C. polystachios* subsp. *atrachus* is very poorly represented, with

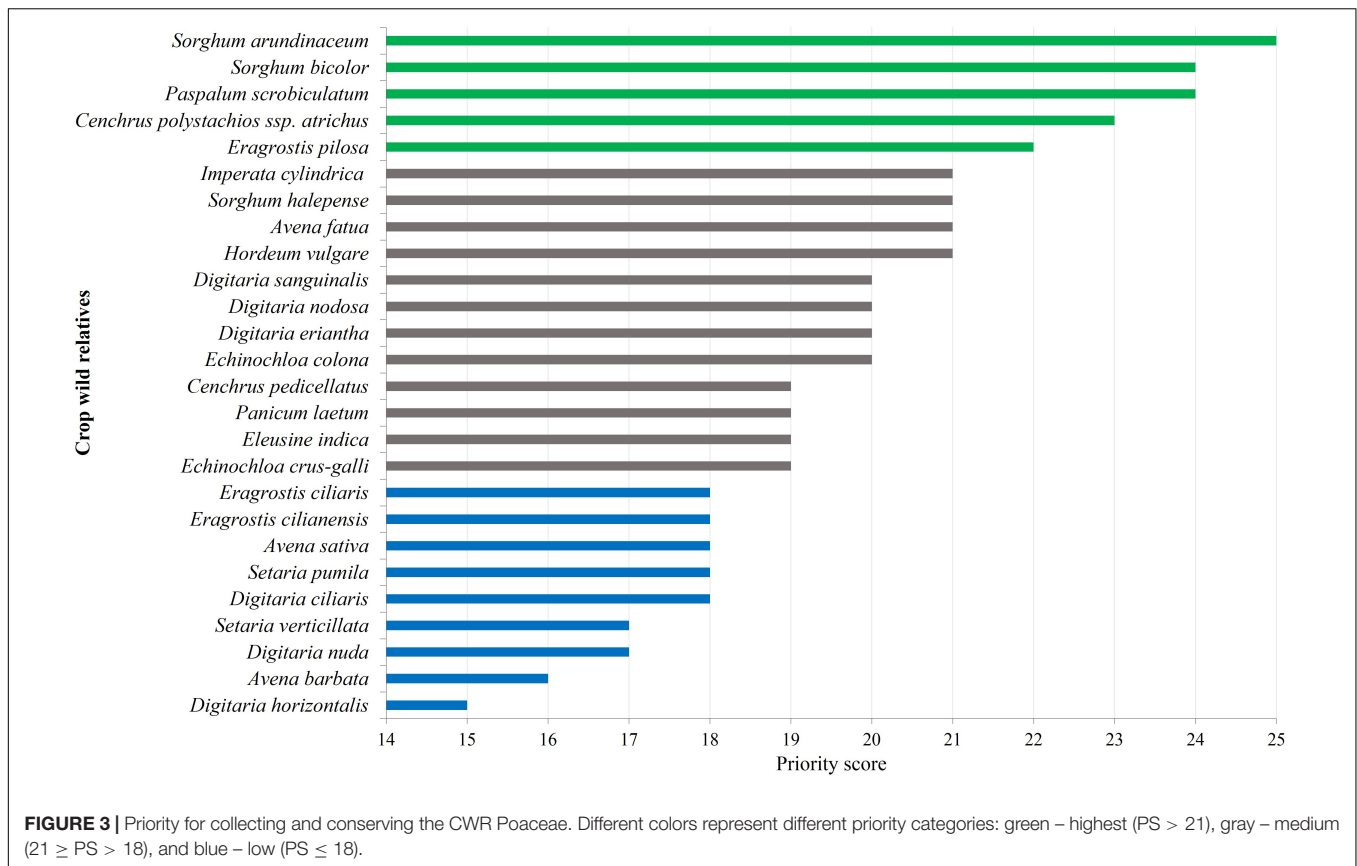
only 3 accessions, two of them from West Africa; and *E. pilosa* has 51 accessions, including nine from West Africa.

Considering only the accessions collected in Cabo Verde, we found one, of *S. bicolor*, hosted in the International Crop Research Institute for the Semi-arid Tropics (ICRISAT) in India.

DISCUSSION

Climate change in Sub-Saharan Africa (SSA) has been impacting water resources, and agricultural and food systems, particularly during the first decade of the 21st century, the warmest decade on record (Rickards and Howden, 2012; Hartmann et al., 2013). Cabo Verde Islands are highly susceptible to climate change due to consecutive years of drought and their poor soil structure, intensified by scarce vegetation cover (Neto et al., 2020).

In drylands, the importance of Poaceae species extends from the cultivation of grasslands and erosion control to, and especially, uses by humans (e.g., supply of cereals) and livestock (e.g., fodder and forage) (Capstaff and Miller, 2018; Mganga et al., 2019). More than half of the population's food supply is provided by three grass crops (i.e., rice, wheat, and maize)



which are particularly important in developing countries where they provide food security and nutrition to local populations (Tadele, 2016).

Our study identified 26 Cabo Verde's CWR from the Poaceae family. This archipelago is an important center of wild diversity of African crop millets, namely of fonio (e.g., white fonio *D. exilis*, and black fonio, *D. iburua*) and other African millets, such as: pearl millet (*C. americanus*), teff millet (*E. tef*), finger millet (*E. coracana*), barnyard millet (*E. colona*), proso millet (*P. miliaceum*), and foxtail millet (*S. italica*). African millets represent a diverse group of cereal crops, which are well adapted to adverse agroecological conditions (Garí, 2002). Millets and their wild forms represent critical plant genetic resources for the agriculture and food security of poor farmers who inhabit arid, infertile, and marginal lands (Teixeira et al., 2013). The Food and Agriculture Organization has announced the year 2023 as 'International Year of Millets', recognizing the potential of these crops to fight malnutrition and hunger in developing countries (Muthamilarasan and Prasad, 2021).

Both millets and sorghum CWR of Cabo Verde occur under extreme climatic conditions in this archipelago, being presumably more resilient to climate change. Some of the CWR species found in Cabo Verde, namely *S. bicolor*, *H. vulgare*, and *A. fatua*, have already shown high tolerance to droughts and saline environments (Dinari et al., 2013; Ogbaga et al., 2014; Gous et al., 2015). Moreover, a study conducted by Gurney et al. (2002) revealed that *Sorghum*

arundinaceum is highly tolerant to infection by *Striga hermonthica* and *Striga asiatica* root hemiparasites known to attack crops and cause great losses in production. Also, *Sorghum halepense* is comparatively less susceptible to downy mildew infection (caused by *Peronosclerospora sorghi*) than the associated crop (Kamala et al., 2002). Thus, Cabo Verde's CWR could be a valuable source of resistance genes to increase the tolerance of their related crops to biotic and abiotic stresses.

According to Adhikari et al. (2015), the estimated yield loss by the end of this century due to climate change is less than 20% for African millets and sorghum, whereas for other grain crops, such as wheat, a reduction of as much as 72% is foreseen, and for rice and soybean, of up to 45%.

In Cabo Verde, the greatest richnesses of CWR were found in Santiago, Santo Antão, and Fogo islands, with ten or more different taxa; the species richness roughly increased with the area and maximum altitude of the island, which is related with more available habitats, in agreement with other studies on endemic flora (Romeiras et al., 2015, 2016). Also, these islands have more agricultural activity but, except for the cultivation of sugarcane, which is the primary irrigated crop of Cabo Verde (Monteiro et al., 2020), there are no official reports on the cultivation/production of the Poaceae crops included in this study. Some CWR are the wild forms of cultivated crops (e.g., *E. colona*, *E. crus-galli*, *H. vulgare*, *P. scrobiculatum*, and *S. bicolor*) and most of them grow on

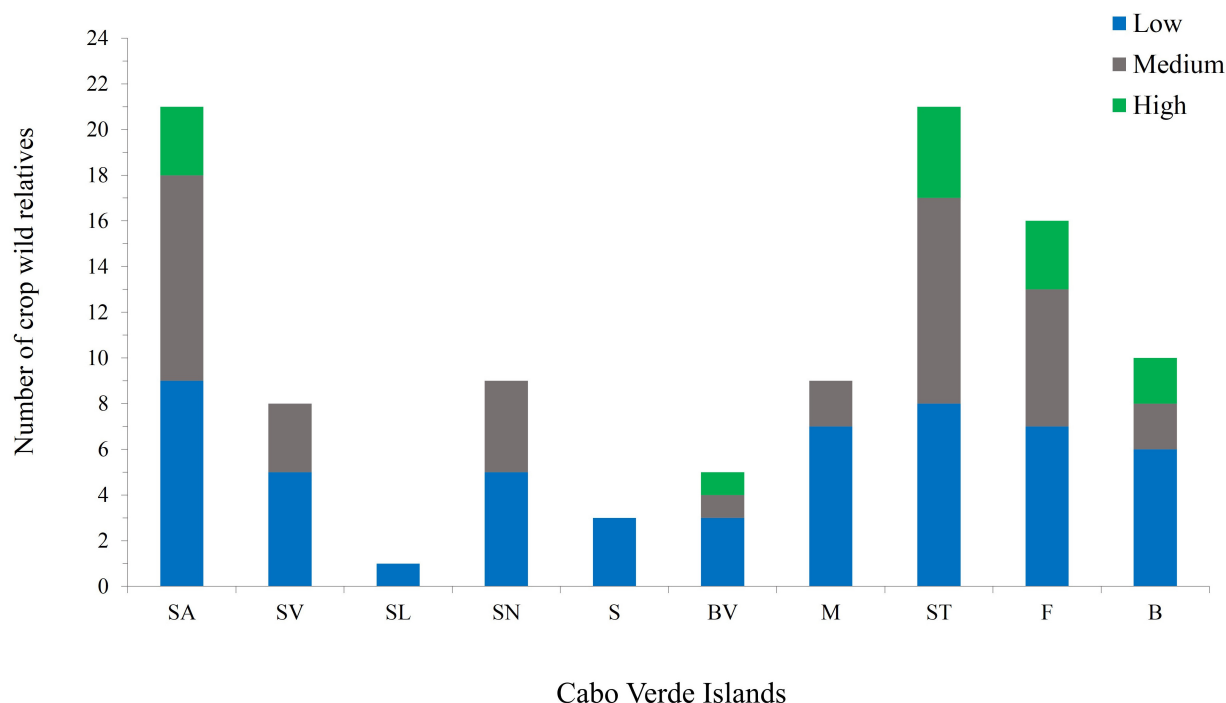


FIGURE 4 | Number of priority CWR Poaceae species per island and their priority categories; total CWR for Cabo Verde archipelago is 26. Different colors represent different priority categories: green – highest; gray – medium; blue – low. Island abbreviations: SA, Santo Antão; SV, São Vicente; SL, Santa Luzia; SN, São Nicolau; S, Sal; BV, Boavista; M, Maio; ST, Santiago; F, Fogo; B, Brava.

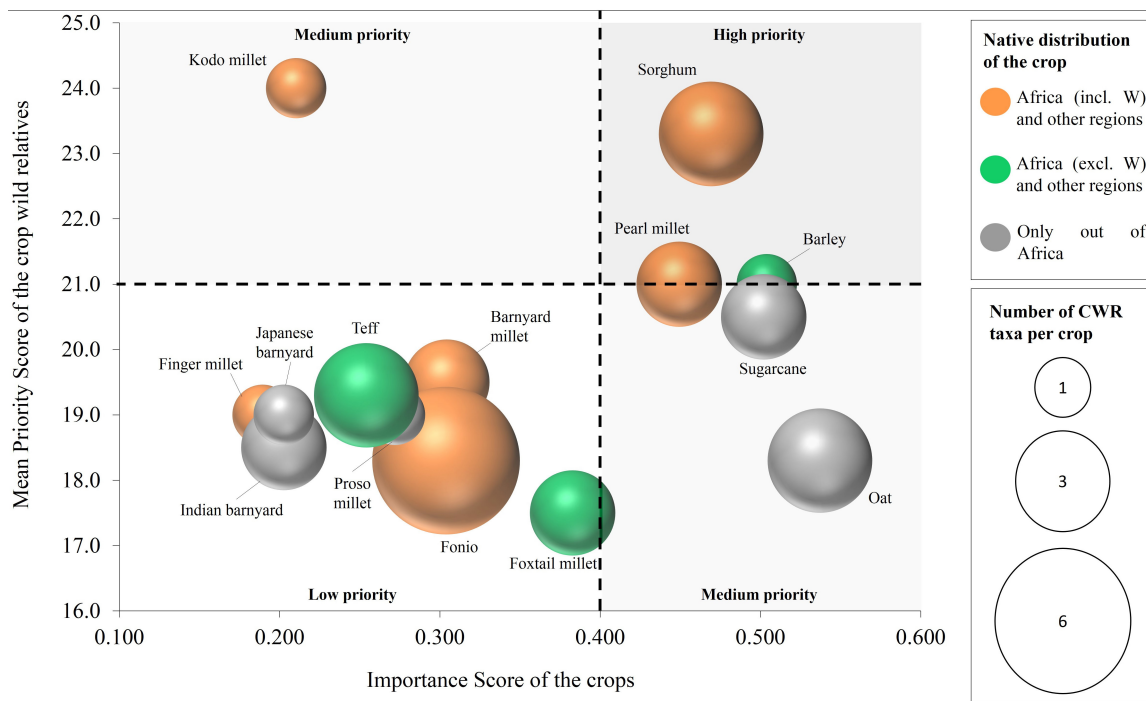


FIGURE 5 | Comparison of the importance of the 14 Poaceae crops studied and their CWR in Cabo Verde. The Importance Score concerns the food supply and agricultural production metrics of the crops and the mean Priority Score represents the nine criteria used as a proxy to prioritize the CWR (for details see section "Importance and Priority Scores"). The size of the circles indicates the number of CWR taxa per crop. The colours indicate the native distribution of the crop.

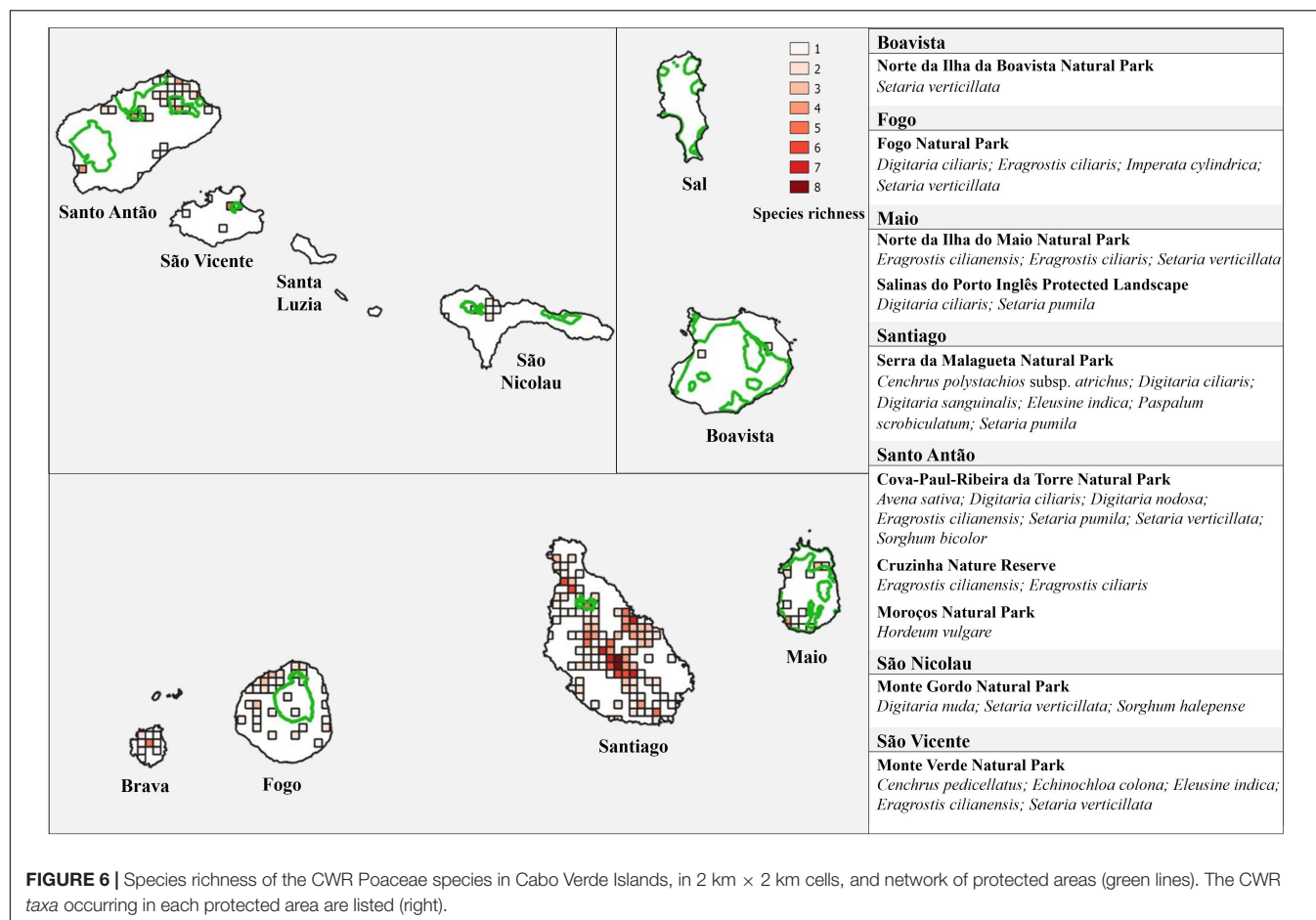


FIGURE 6 | Species richness of the CWR Poaceae species in Cabo Verde Islands, in 2 km × 2 km cells, and network of protected areas (green lines). The CWR taxa occurring in each protected area are listed (right).

marginal agricultural areas, mainly of maize, and on grazing areas. Barnyard millet (*E. colona*) occurs in lowland areas of Santo Antão, São Vicente, São Nicolau, Maio and Santiago; it is well adapted to arid ecosystems and tolerates soils with poor fertility. Nevertheless, *H. vulgare* and *I. cylindrica*, the CWR of sugarcane (*S. officinarum*) stand out for their ability to develop at high altitudes (averaging 1,306 and 1,560 m, respectively). Currently, sugarcane and maize are the most cultivated grass crops in Cabo Verde, but other species are used in their wild forms, for human consumption and, particularly, as forage (Monteiro et al., 2020). The cultivated wild forms of millets could offer a more sustainable food source than their related major crops because they are more efficient in the use of water and nitrogen (Muthamilarasan and Prasad, 2021).

Among the Poaceae species ranked as of high priority for further collection in Cabo Verde, there are *C. polystachios* subsp. *atrichus*, *P. scrobiculatum*, *S. arundinaceum*, *S. bicolor*, and CWRs associated with fonio; all of these species are used in the archipelago for human and animal consumption, as well as for medicinal purposes (Romeiras et al., 2011).

In the past, these food plants have played an important role in the diet and traditional medicine of African communities (Catarino et al., 2016; Havik et al., 2018; Akinola et al., 2020). Several of these species are important African crops (Amadou

et al., 2013; Tadele, 2016), but in Cabo Verde there are no evidences of their cultivation, most of them being used for livestock grazing (Barbosa, 1961) or has building materials (e.g., thatching with *I. cylindrica*).

Some of the reported Poaceae species are used since the colonization of Cabo Verde (Barbosa, 1961; Romeiras et al., 2011). Since the mid-1500s, Cabo Verde has become the subject of numerous travel descriptions reporting on the local flora and on the introduction of new crops such as sugarcane, maize and cotton (Romeiras et al., 2018). Other reports, by the chronicler Valentim Fernandes, confirmed the large production and abundance of sorghum and pearl millet in the West African region, from where they were imported at very early stages of the colonization of this archipelago (Romeiras et al., 2014). These African millets, together with rice, were the most important crops during the settlement of Cabo Verde and the basis of the population's diet (Santos and Torrão, 1998). In the 17th century, pearl millet (*milho zaburro*, *milho de maçaroca* or *milho branco*, as was locally called), began to be intensively cultivated and made this archipelago self-sufficient in cereals, only needing to import cereals from West Africa in drought years (Teixeira da Mota and Carreira, 1966). The introduction and cultivation of maize (*Zea mays* L.) gradually replaced the role that African millets had until the 18th century (Torrão, 1995). Ancient reports refer

TABLE 4 | Accessions of the studied species available in international Germplasm Banks.

CWR occurring in Cabo Verde	Provenance and number of accessions			Countries holding accessions	Total accessions
	West Africa	Africa (except West Africa)	Other Regions		
<i>Avena barbata</i>	0	240 [Morocco (199); Libya (41)]	988	> 10	1228
<i>Avena fatua</i>	0	0	1876	> 10	1876
<i>Avena sativa</i> *	0	0	53617	> 10	53617
<i>Cenchrus pedicellatus</i>	30 [Niger (17); Mali (9); Nigeria (2); Burkina Faso (1); Mauritania (1)]	62 [Cameroon (56); Central African Republic (3); Ethiopia (3)]	61	5	153
<i>Cenchrus polystachios</i> subsp. <i>atrichus</i>	2 [Burkina Faso (1); Mali (1)]	1 [Central African Republic (1)]	0	1	3
<i>Digitaria ciliaris</i>	1 [Niger (1)]	8 [Tanzania (3); Mozambique (2); Kenya (1); Madagascar (1); Republic of South Africa (1)]	9	5	18
<i>Digitaria eriantha</i>	0	584	113	6	697
<i>Digitaria horizontalis</i>	6 [Burkina Faso (5); Nigeria (1)]	0	1	4	7
<i>Digitaria nodosa</i>	0	4 [Kenya (4)]	1	2	5
<i>Digitaria nuda</i>	0	0	0	0	0
<i>Digitaria sanguinalis</i>	0	2 [Malawi (2)]	32	> 10	34
<i>Echinochloa colona</i> *	19 [Mali (19)]	21 [Kenya (8); Republic of South Africa (4); Botswana (3); Ethiopia (2); Malawi (2); Sudan (2)]	520	7	660
<i>Echinochloa crus-galli</i> *	0	0	350	10	350
<i>Eleusine indica</i>	10 [Nigeria (7); Ghana (3)]	76 [Uganda (48); Kenya (22); Democratic Republic of Congo (4); Burundi (2)]	82	> 10	168
<i>Eragrostis cilianensis</i>	9 [Mali (6); Burkina Faso (3)]	13 [Kenya (8); Madagascar (4); Ethiopia (1)]	12	6	34
<i>Eragrostis ciliaris</i>	3 [Mali (2); Burkina Faso (1)]	9 [Kenya (8); Madagascar (1)]	4	2	16
<i>Eragrostis pilosa</i>	9 [Burkina Faso (6); Mali (3)]	23 [Kenya (18); Ethiopia (3); Madagascar (2)]	19	6	51
<i>Hordeum vulgare</i> *	0	18391 [Ethiopia (all)]	234154	> 10	252545
<i>Imperata cylindrica</i>	0	9 [Madagascar (3); Kenya (2); Tanzania (2); Malawi (1); Republic of South Africa (1)]	12	4	21
<i>Panicum laetum</i>	28 [Mali (24); Niger (2); Mauritania (1); Burkina Faso (1)]	0	0	2	28
<i>Paspalum scrobiculatum</i> *	13 [Mali (9); Niger (4)]	79 [Kenya (29); Zimbabwe (17); Tanzania (7); Uganda (7); Madagascar (5); Republic of South Africa (5); Ethiopia (9)]	1022	7	1114
<i>Setaria pumila</i>	0	6 [Cameroon (4); Kenya (2)]	81	10	87
<i>Setaria verticillata</i>	1 [Burkina Faso (1)]	4 [Botswana (4)]	28	9	33
<i>Sorghum arundinaceum</i>	11 [Mali (11)]	311 [Sudan (145); Kenya (42); Republic of South Africa (34); Uganda (30); Angola (19); Ethiopia (18); Egypt (13); Chad (10)]	108	9	430
<i>Sorghum bicolor</i> *	6729 [Nigeria (3488); Mali (3240); Cabo Verde (1; Accession number – IS 27941; PGRFA doi: 10.18730/NS7TT)]	37265 [Ethiopia (12112); Sudan (11444); Kenya (6526); Zimbabwe (4152); Uganda (3031)]	79479	> 10	123473
<i>Sorghum halepense</i>	0	17 [Sudan (9); Angola (4); Republic of South Africa (4)]	186	> 10	203

CWR taxon: * = The wild form of the crop. Due to the floristic affinities with Cabo Verde and West African floras, the provenance of the accessions is provided. Accession number and DOI are included to the only Cabo Verde accession.

that maize was known by the Portuguese since the 16th century, due to regular trade between Cabo Verde and the Antilles, where this species is native. However, the cultivation of African millets, more adapted to poor soils and with less water requirements, was more appropriate than that of the American maize (Santos and Torrão, 1998). Since the end of the journeys to the Antilles and the establishment of regular trade with Brazil, maize became the food basis of the Cabo Verdean population (Santos and Torrão, 1998) and is currently a dominant part of their diet (Monteiro et al., 2020). The importance of historical factors and the role played by Cabo Verde in the Atlantic navigation during the 16th–19th centuries (Romeiras et al., 2020), contributed to change and determine the present composition of Cabo Verde's flora, with more than 70% of exotic species, most of them introduced for food purposes.

Therefore, the valorisation of the plant genetic resources related to millet crops and their wild relatives is of major importance to fight hunger and ensure food and nutrition security in Cabo Verde. This archipelago is still very dependent on food importations, particularly in years of prolonged and severe droughts, as happened in 2017/2018 (Monteiro et al., 2020). Adding to the adverse natural conditions, it is a small and fragmented insular country, with inherent difficulties in connections with West Africa and Europe, as well as between islands, posing problems to the rapid provision of food (either locally produced or imported) to more inaccessible rural areas, namely in Brava, São Nicolau, and Maio Islands. Despite recent progress in reducing extreme poverty, ca. 30% of Cabo Verde's population still lives in multidimensional poverty with poor health care, lack of education, and inadequate living standards, which is a strong economic obstacle to meet the food needs of a large section of the population (Varela et al., 2020).

Notwithstanding the recognized importance of African millets, information available on online databases, such as GRIN (USDA, 2020), revealed that data on these species and their wild relatives are still limited. Also, the worldwide CWR inventory¹ revealed that there is no information about confirmed traits for millets. Therefore, *ex situ* conservation of plant genetic resources of Cabo Verde must be a national priority in response to the rapid loss of agricultural biodiversity, as there is only one accession (*S. bicolor*, Genesys, 2020) available for Cabo Verdean Poaceae CWRs in genebanks. New expeditions must be performed in these islands, to collect CWR species growing in threatened habitats, and targeted to priority species with few accessions stored in worldwide genebanks, such as *C. polystachios* subsp. *atrichus*, *E. pilosa*, *I. cylindrica*, *P. laetum*, and *S. arundinaceum*. Most of these taxa have native distribution ranges in West Africa and are important for human consumption and to feed livestock. Moreover, and although their use is less widespread, some species of the genus *Urochloa* (= *Brachiaria*) present in Cabo Verde are considered small millets in Africa. Such is the case of *Urochloa deflexa* (Schumach.) H.Scholz [= *Brachiaria deflexa* (Schumach.) C.E.Hubb. ex Robyns] (guinea millet), used as food in the Sudan-Zambezi and Yemenite regions in its wild form and, occasionally, cultivated in the highlands of the Fouta Djallon (Portères, 1976),

and of *Urochloa ramosa* (L.) T.Q.Nguyen [= *Brachiaria ramosa* (L.) Stapf] (browntop millet) more widespread as forage, but also used as food in India (Kimata et al., 2000). In Cabo Verde, these species are mostly referred to as forage. There is no record of genetic material from Cabo Verde, where one of the five existing *Urochloa* species is endemic [= *Urochloa caboverdiana* (Conert & C.Kohler) Veldkamp, Potdar & S.R.Yadav], reinforcing the need for its *ex situ* conservation.

Although *ex situ* conservation has had more worldwide success than *in situ* conservation, probably because of its facility of access by users and lower cost (De-Zhu and Pritchard, 2009; Díez et al., 2018), the establishment of the Protected Areas Network in Cabo Verde has already contributed to safeguard the archipelago's natural heritage and endemic species (MAAP, 2004; Romeiras et al., 2016). Our study was able to identify hotspot areas for *in situ* conservation of CWR populations across the archipelago, and species were identified and correlated with habitat conditions, namely to detect which ones are better adapted to drylands, highlands, and poor soils in these islands.

FINAL REMARKS

The benefits of grass crops, namely African millets, and their ancestral use in Cabo Verde, were highlighted in our study, which also alerts to the need of rescuing cultural values, and to the consumer's unawareness of the advantages of these plants, well adapted to the very dry conditions of this archipelago. However, the adverse natural environmental conditions of a small and fragmented insular country such as Cabo Verde, with inherent difficulties in inter-island transportation, hinder the supply of food products to rural populations. So, it is necessary to produce more and with better quality under various limitations, such as marginal lands, water shortage, soil degradation, or climate change (Castañeda-Álvarez et al., 2016). In this context, viable approaches to improve food security are crucial, and the systematic use of CWR in crop improvement appears essential to face the increasing pressure on food production while maintaining natural diversity. Cabo Verde's plant diversity faces increasing threats (Romeiras et al., 2016) due to desertification processes, and native species remain a viable sustainable land management option to fight degradation in these tropical dry islands.

DATA AVAILABILITY STATEMENT

The original contributions generated for this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MD and MR: conceptualization and supervision. VR, MD, SC, and MR: methodology and writing—review and editing.

¹<https://www.cwrdiversity.org/checklist/>

VR and SC: formal analysis. MD, VR, and ID: field surveys. VR: investigation. MR: writing—original draft preparation. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the Fundação para a Ciência e Tecnologia (FCT) and Aga Khan Development Network (AKDN) through the project CVAgrobiodiversity/333111699. VR was supported by FCT/AGAKAN grant and SC by FCT grant (SFRH/BD/120054/2016).

REFERENCES

- Adhikari, U., Nejadhashemi, A. P., and Woznicki, S. A. (2015). Climate change and eastern Africa: a review of impact on major crops. *Food Energy Sec.* 4, 110–132. doi: 10.1002/fes3.61
- Akinola, R., Pereira, L. M., Mabhaudhi, T., de Bruin, F. M., and Rusch, L. (2020). A review of indigenous food crops in africa and the implications for more sustainable and healthy food systems. *Sustainability* 12:3493. doi: 10.3390/su12083493
- Allen, E., Gaisberger, H., Magos Brehm, J., Maxted, N., Thormann, I., Lupupa, T., et al. (2019). A crop wild relative inventory for Southern Africa: a first step in linking conservation and use of valuable wild populations for enhancing food security. *Plant Gen. Resour.* 17, 1–12. doi: 10.1017/S1479262118000515
- Amadou, I., Gounga, M. E., and Le, G.-W. (2013). Millets: nutritional composition, some health benefits and processing – a review. *Food Sci. Nutr.* 25, 501–508. doi: 10.9755/efja.v25i7.12045
- Arechavaleta, M., Zurita, N., Marrero, M. C., and Martín, J. L. (eds) (2005). *Lista Preliminar de Espécies Silvestres de Cabo Verde (Hongos, Plantas y Animales Terrestres)*. Santa Cruz de Tenerife: Consejería de Medio Ambiente y Ordenación Territorial, 155.
- Barbosa, L. A. G. (1961). Subsídios para um dicionário utilitário e glossário dos nomes vernáculos das plantas de arquipélago de Cabo Verde. *Garcia Orta Sér. Bot.* 9, 37–91.
- Belton, P. S., and Taylor, J. R. N. (2004). Sorghum and millets: protein sources for Africa. *Trends Food Sci. Technol.* 15, 94–98. doi: 10.1016/j.tifs.2003.09.002
- Capstaff, N. M., and Miller, A. J. (2018). Improving the yield and nutritional quality of forage crops. *Front. Plant Sci.* 9:535. doi: 10.3389/fpls.2018.00535
- Castañeda-Álvarez, N. P., Khoury, C. K., Achicanoy, H. A., Bernau, V., Dempewolf, H., Eastwood, R. J., et al. (2016). Global conservation priorities for crop wild relatives. *Nat. Plants* 2:16022. doi: 10.1038/nplants.2016.22
- Catarino, L., Havik, P., and Romeiras, M. M. (2016). Medicinal plants of Guinea-Bissau: therapeutic applications, ethnic diversity and knowledge transfer. *J. Ethnopharmacol.* 183, 71–94. doi: 10.1016/j.jep.2016.02.032
- CGIAR-CSI Consortium for Spatial Information (2020). *SRTM 90m Digital Elevation Database v4.1*. Available online at: <https://cgiasi.community/data/srtm-90m-digital-elevation-database-v4-1/> (accessed on 16 June 2020).
- Chandi, G. K., and Annor, G. A. (2016). *Millet Minor: Overview*. In: *Encyclopedia of Food Grains*. Amsterdam: Elsevier Inc., 199–208.
- Chapman, A. D., and Wiczorek, J. (2006). *Guide to Best Practices for Georeferencing*. Copenhagen: Global Biodiversity Information Facility.
- CWR (2019). *The Crop Wild Relative Project*. Available online at: <https://www.cwrdiversity.org/checklist/> (accessed on 15 March 2020)
- Dempewolf, H., Eastwood, R. J., Guarino, L., Khoury, C. K., Müller, J. V., and Toll, J. (2014). Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecol. Sustain. Food Syst.* 38, 369–377. doi: 10.1080/21683565.2013.870629
- De-Zhu, L., and Pritchard, H. W. (2009). The science and economics of ex situ plant conservation. *Trends Plant Sci.* 14, 614–621. doi: 10.1016/j.tplants.2009.09.005

ACKNOWLEDGMENTS

We thank Fundação para a Ciência e Tecnologia (FCT) and Aga Khan Development Network (AKDN) for funding the project CVAgrobiodiversity. We also thank the research units: UID/AGR/04129/2020 (LEAF) and UID/BIA/00329/2020 (cE3c) funded by Portuguese National Funds through FCT, Portugal.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.630217/full#supplementary-material>

- Díez, M. J., De la Rosa, L., Martín, I., Guasch, L., Carrea, M. E., Mallor, C., et al. (2018). Plant genebanks: present situation and proposals for their improvement. The case of the Spanish Network. *Front. Plant Sci.* 9:1794. doi: 10.3389/fpls.2018.01794
- Dinari, A., Meighani, F., and Farzami, S. M. (2013). Effects of salinity and drought stress on germination and seedling growth of *Avena fatua* L. and *Phalaris minor* L. *Iran. J. Plant Physiol.* 3, 665–671.
- Duarte, M. C., Rego, F., Romeiras, M. M., and Moreira, I. (2008). Plant species richness in the Cape Verde Islands – eco-geographical determinants. *Biodivers. Conserv.* 17, 453–466. doi: 10.1007/s10531-007-9226-y
- Duarte, M. C., and Romeiras, M. M. (2009). “Cape Verde Islands,” in *Encyclopedia of Islands*, eds R. Gillespie and D. Clague (Berkeley: University of California Press), 143–148. doi: 10.1525/9780520943728-033
- Dwivedi, S., Upadhyaya, H., Senthilvel, S., Hash, C., Fukunaga, K., Diao, X., et al. (2012). Millets: genetic and genomic resources. *Plant Breed. Rev.* 35, 247–375. doi: 10.1002/9781118100509.ch5
- Essoh, A. P., Monteiro, F., Pena, A. R., Pais, M. S., Moura, M., and Romeiras, M. M. (2020). Exploring glucosinolates diversity in Brassicaceae: a genomic and chemical assessment for deciphering abiotic stress tolerance. *Plant Physiol. Biochem.* 150, 151–161. doi: 10.1016/j.plaphy.2020.02.032
- Food and Agricultural Organization of the United Nations [FAOSTAT] (2020). *Food and Agricultural Data*. Available online at: <http://www.fao.org/faostat/en/#home> (accessed on 27 April 2020)
- Ford-Lloyd, B. V., Schmidt, M., Armstrong, S. J., Barazani, O., Engels, J., Hadas, R., et al. (2011). Crop wild relatives – undervalued, underutilized and under threat? *BioScience* 61, 559–565. doi: 10.1525/bio.2011.61.7.10
- Freitas, R., Romeiras, M., Silva, L., Cordeiro, R., Madeira, P., González, J. A., et al. (2019). Restructuring of the ‘Macaronesia’ Biogeographic unit: a marine multitaxon biogeographical approach. *Sci. Rep.* 9:15792. doi: 10.1038/s41598-019-51786-6
- Gari, J. A. (2002). “Review of the African millet diversity,” in *Proceedings of the International Workshop on Fonio, Food Security and Livelihood Among the Rural Poor in West Africa*, (Bamako: IPGRI/IFAD), 19–22.
- GBIF.org (2020). *GBIF Occurrence Download*. Available online at: <https://doi.org/10.15468/dl.9gm3g3> (accessed on 30 June 2020).
- Genesys (2020). *Plant Genetic Resources Accession*. Available online at: <https://www.genesys-pgr.org/> (accessed 19 July 2020)
- Gous, P. W., Gilbert, R. G., and Fox, G. P. (2015). Drought-proofing barley (*Hordeum vulgare*) and its impact on grain quality: a review. *J. Inst. Brew.* 121, 19–27. doi: 10.1002/jib.187
- Gurney, A. L., Press, M. C., and Scholes, J. D. (2002). Can wild relatives of sorghum provide new sources of resistance or tolerance against *Striga* species? *Weed Res.* 42, 317–324. doi: 10.1046/j.1365-3180.2002.00291.x
- Harlan, J. R., and de Wet, J. M. J. (1971). Towards a rational classification of cultivated plants. *Taxon* 20, 509–517. doi: 10.2307/1218252
- Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L. V., Brönnimann, S., Charabi, Y., et al. (eds) (2013). “Observations: atmosphere and surface,” in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel*

- on *Climate Change*, eds T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Cambridge: Cambridge University Press).
- Havik, P. J., Monteiro, F., Catarino, S., Correia, A. M., Catarino, L., and Romeiras, M. M. (2018). Agro-economic transitions in guinea-bissau (West Africa): historical trends and current insights. *Sustainability* 10:3408. doi: 10.3390/su10103408
- Heywood, V., Casas, A., Ford-Lloyd, B., Kell, S., and Maxted, N. (2007). Conservation and sustainable use of crop wild relatives. *Agric. Ecosyst. Environ.* 121, 245–255. doi: 10.1016/j.agee.2006.12.014
- ICRISAT (2020). ICRISAT. Available online at: <https://www.icrisat.org/> (accessed on 9 April 2020)
- IDE-CV (2020). *Infra-Estrutura de Dados Espaciais de Cabo Verde*. Available online at: <https://idecv-ingt.opendata.arcgis.com/> (accessed on 06 October 2020).
- Ingram, A. L., and Doyle, J. J. (2007). *Eragrostis* (Poaceae): monophyly and infrageneric classification. *Aliso J. Syst. Evol. Bot.* 23, 595–604. doi: 10.5642/aliso.20072301.44
- IPGRI (2002). *Programme for Neglected and Underutilised Species*. Rome: International Plant Genetic Resources Institute.
- Kajuna, S. T. A. R. (2001). "Millet: post-harvested operations," in *FAO INPhO – Post-Harvest Compendium*, eds D. Mejía and B. Lewis (Morogoro: Sokone University of Agriculture).
- Kamala, V., Singh, S. D., Bramel, P. J., and Rao, D. M. (2002). Sources of resistance to downy mildew in wild and weedy sorghums. *Crop Sci.* 42, 1357–1360. doi: 10.2135/cropsci2002.1357
- Kimata, M., and Ashok, E. G., and Seetharam, A. (2000). Domestication, cultivation and utilization of two small millets, *Brachiaria ramosa* and *Setaria glauca*, Poaceae in South India. *Econ. Bot.* 54, 217–227. doi: 10.1515/9783110806373.409
- Lobin, W. (1986). Katalog der von den Kapverdischen Inseln beschriebenen Taxa höherer Pflanzen (Pteridophyta & Phanerogamae). *Cour. Forschungsinst. Senckenberg* 81, 93–164.
- MAAP (2004). *Livro Branco sobre o Estado do Ambiente em Cabo Verde*. North Melbourne VIC: MAAP.
- Maxted, N., Ford-Lloyd, B. V., Jury, S., Kell, S., and Scholten, M. (2006). Towards a definition of a crop wild relative. *Biodivers. Conserv.* 15, 2673–2685. doi: 10.1007/s10531-005-5409-6
- Maxted, N., Magos Brehm, J., and Kell, S. (2013). *Resource Book for Preparation of National Conservation Plans for Crop Wild Relatives and Landraces*. Birmingham: University of Birmingham.
- Mganga, K. Z., Nyariki, D. M., Musimba, N. K., and Mwang'ombe, A. W. (2019). "Indigenous grasses for rehabilitating degraded African drylands," in *Agriculture and Ecosystem Resilience in Sub Saharan Africa*, eds Y. Bamatuze, S. Kyamanywa, B. R. Singh, G. Nabanoga, and R. Lal (Cham: Springer), 53–68.
- Monteiro, F., Fortes, A., Ferreira, V., Essoh, A. P., Gomes, I., Correia, A. M., et al. (2020). Current status and trends in Cabo Verde agriculture. *Agronomy* 10:74. doi: 10.3390/agronomy10010074
- Muthamilarasan, M., and Prasad, M. (2021). Small millets for enduring food security amidst pandemics. *Trends Plant Sci.* 26, 33–40. doi: 10.1016/j.tplants.2020.08.008
- Neto, C., Costa, J. C., Figueiredo, A., Capelo, J., Gomes, I., Vitória, S., et al. (2020). The role of climate and topography in shaping the diversity of plant communities in Cabo Verde Islands. *Diversity* 12:80. doi: 10.3390/d12020080
- Ogbaga, C. C., Stepien, P., and Johnson, G. N. (2014). Sorghum (*Sorghum bicolor*) varieties adopt strongly contrasting strategies in response to drought. *Physiol. Plant.* 152, 389–401. doi: 10.1111/ppl.12196
- Portères, R. (1976). "African Cereals: Eleusine, Fonio, Black Fonio, Tejf, *Brachiaria*, *Paspalum*, *Pennisetum*, and African Rice," in *Origins of African Plant Domestication*, ed. J. R. Harlan (Mouton: The Hague), 409–452.
- POWO (2019). *Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew*. Available online at: <http://www.plantsoftheworldonline.org/> (accessed on 15 January 2020)
- PROTA4U (2020). *Plant Resources of Tropical Africa*. Available online at: <https://www.prota4u.org/database/> (accessed on 25 March 2020)
- QGIS Development Team (2020). *QGIS Geographic Information System. Open Source Geospatial Foundation Project*. Available online at: <http://qgis.osgeo.org> (accessed October 6, 2020).
- R Development Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rickards, L., and Howden, S. M. (2012). Transformational adaptation: agriculture and climate change. *Crop Past. Sci.* 63, 240–250.
- Rivas-Martínez, S., Lousã, M., Costa, J. C., and Duarte, M. C. (2017). Geobotanical survey of the Cabo Verde Islands (West Africa). *Int. J. Geobot. Res.* 7, 1–103.
- Romeiras, M. M., Carine, M., Duarte, M. C., Catarino, S., Dias, F. S., and Borda-de-Água, L. (2020). Bayesian methods to analyze historical collections in time and space: a case study using Cabo Verde Endemic Flora. *Front. Plant Sci.* 11:278. doi: 10.3389/fpls.2020.00278
- Romeiras, M. M., Catarino, L., Torrão, M. M., and Duarte, M. C. (2011). Diversity and origin of medicinal exotic flora in Cape Verde Islands. *Plant Ecol. Evol.* 142, 214–225. doi: 10.5091/pleveo.2011.560
- Romeiras, M. M., Catarino, S., Filipe, A., Magalhães, M., Duarte, M. C., and Beja, P. (2016). Species conservation assessments in small islands: the consequences of precautionary versus evidentiary attitudes. *Conserv. Lett.* 9, 275–280. doi: 10.1111/conl.12212
- Romeiras, M. M., Catarino, S., Gomes, I., Fernandes, C., Costa, J. C., Caujapé-Castells, J., et al. (2015). IUCN red list assessment of the Cape Verde endemic flora: towards a Global strategy for plant conservation in Macaronesia. *Bot. J. Linnean Soc.* 180, 413–425. doi: 10.1111/boj.12370
- Romeiras, M. M., Duarte, M., Francisco-Ortega, J., Catarino, L., and Havik, P. (2018). Recovering plant data for guinea-bissau: implications for biodiversity knowledge of West Africa. *Diversity* 10:109. doi: 10.3390/d10040109
- Romeiras, M. M., Duarte, M. C., Santos-Guerra, A., Carine, M., and Francisco-Ortega, J. (2014). Botanical exploration of the Cape Verde Islands: from the pre-Linnaean records and collections to late 18th century floristic accounts and expeditions. *Taxon* 63, 625–640. doi: 10.12705/633.37
- Santos, M. E. M., and Torrão, M. M. (1998). "Entre l'Amérique et l'Afrique, les îles du Cap-Vert et São Tomé: Les cheminements des milhos (mil, sorgho et maïs)," in *Plantes et paysages d'Afrique: une histoire à explorer*, ed. M. Chastanet (Paris: Karthala Editions).
- Serea, R. (2018). *Google Earth Pro* 7, 3.2.5491. Available online at: <https://www.neowin.net/news/google-earth-pro-7325491/> (accessed on 06 March 2020)
- Sonesson, C., Davidson, G., and Sachs, L. (2016). *Mapping Mining to the Sustainable Development Goals: A Preliminary Atlas*. Cologny: World Economic Forum.
- Tadele, Z. (2016). "Drought adaptation in millets. Chapter 26," in *Abiotic and Biotic Stress in Plants - Recent Advances and Future Perspectives*, eds A. K. Shanker and C. Shanker (Rijeka: IntechOpen).
- Taylor, J. R. N. (2016). Millet pearl: overview. *Encycl. Food Grains* 1, 190–198. doi: 10.1016/B978-0-12-394437-5.00011-5
- Teixeira, A. J. S., and Barbosa, L. A. G. (1958). "A agricultura do arquipélago de Cabo Verde: Cartas agrícolas, problemas agrários," in *Memórias da Junta de Investigações do Ultramar*, 2, 178.
- Teixeira, da Mota, A., and Carreira, A. (1966). "Milho Zaburro" and "Milho Maçaroca" in Guinea and in the Islands of Cabo Verde. *J. Int. Afric. Inst.* 36, 73–84.
- Teixeira, E. I., Fischer, G., van Velthuisen, H., Walter, C., et al. (2013). Global hot-spots of heat stress on agricultural crops due to climate change. *Agric. For. Meteorol.* 170, 206–215.
- Teso, M. L. R., Lamas, E. T., Parra-Quijano, M., de la Rosa, L., Fajardo, J., and Iriondo, J. M. (2018). National inventory and prioritization of crop wild relatives in Spain. *Genet. Resour. Crop. Evol.* 65, 1237–1253. doi: 10.1007/s10722-018-0610-0
- The Plant List (2013). *Version 1.1*. Available online at: www.theplantlist.org/ (accessed on 25 September 2019)
- Torrão, M. M. (1995). *Diets Alimentares. Transferências e Adaptações nas Ilhas de Cabo Verde (1460–1540)*. Lisboa: Instituto de Investigação Científica Tropical, 108.
- United Nations [UN] (2019). *Department of Economic and Social Affairs, Population Division. World Population Prospects 2019*. Available online at: <http://esa.un.org/unpd/wpp/index.htm> (accessed on 2 May 2020)
- USDA (2020). "Agricultural research service, and national plant germplasm system," in *Germplasm Resources Information Network (GRIN-Taxonomy)*. Beltsville, MD: National Germplasm Resources Laboratory.
- van Treuren, R., Hoekstra, R., and van Hintum, T. J. L. (2017). Inventory and prioritization for the conservation of crop wild relatives in The Netherlands

- under climate change. *Biol. Conserv.* 216, 123–139. doi: 10.1016/j.biocon.2017.10.003
- Varela, D., Monteiro, F., Vidigal, P., Silva, L., and Romeiras, M. M. (2020). Mechanisms implemented for the sustainable development of agriculture: an overview of Cabo Verde performance. *Sustainability* 12:5855. doi: 10.3390/su12145855
- Varshney, R., Shi, C., Thudi, M., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976. doi: 10.1038/nbt.3943
- Vetriventhan, M., Azevedo, V. C. R., Upadhyaya, H. D., et al. (2020). Genetic and genomic resources, and breeding for accelerating improvement of small millets: current status and future interventions. *Nucleus* 63, 217–239. doi: 10.1007/s13237-020-00322-3
- Vincent, H., Wiersema, J., Kell, S., Fielder, H., Dobbie, S., Castañeda-Álvarez, N. P., et al. (2013). A prioritized crop wild relative inventory to help underpin global food security. *Biol. Conserv.* 167, 265–275. doi: 10.1016/j.biocon.2013.08.011
- Williams, I. O., Onyenweaku, E. O., and Atangwho, I. J. (2016). Nutritional and antimicrobial evaluation of *Saccharum officinarum* consumed in Calabar, Nigeria. *Afric. J. Biotechnol.* 15, 1789–1795. doi: 10.5897/AJB2015.14877
- Wrigley, C. W. (2016). An overview of the family of cereal grains prominent in the world agriculture. *Encycl. Food Grains* 1, 73–85. doi: 10.1016/B978-0-12-394437-5.00006-1
- Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., and Song, B. H. (2017). Back into the wild – apply untapped genetic diversity of wild relatives for crop improvement. *Evol. Appl.* 10, 5–24. doi: 10.1111/eva.12434

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Rocha, Duarte, Catarino, Duarte and Romeiras. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Genetic Variation for Nitrogen Use Efficiency Traits in Global Diversity Panel and Parents of Mapping Populations in Pearl Millet

Vijayalakshmi Pujarula, Madhu Pusuluri, Srikanth Bollam, Roma Rani Das, Rambabu Ratnala, Gopikrishna Adapala, Vishnukiran Thuraga, Abhishek Rathore, Rakesh K. Srivastava* and Rajeev Gupta*

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India

OPEN ACCESS

Edited by:

Velu Govindan,
International Maize and Wheat
Improvement Center, Mexico

Reviewed by:

Chandirakala Ramakrishnan,
Tamil Nadu Agricultural University,
India

Mehanathan Muthamilarasan,
University of Hyderabad, India

*Correspondence:

Rakesh K. Srivastava
r.k.srivastava@cgiar.org
Rajeev Gupta
g.rajeev@cgiar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 04 November 2020

Accepted: 11 January 2021

Published: 04 February 2021

Citation:

Pujarula V, Pusuluri M, Bollam S, Das RR, Ratnala R, Adapala G, Thuraga V, Rathore A, Srivastava RK and Gupta R (2021) Genetic Variation for Nitrogen Use Efficiency Traits in Global Diversity Panel and Parents of Mapping Populations in Pearl Millet. *Front. Plant Sci.* 12:625915. doi: 10.3389/fpls.2021.625915

Nitrogen (N) is one of the primary macronutrients required for crop growth and yield. This nutrient is especially limiting in the dry and low fertility soils where pearl millet [*Pennisetum glaucum* (L.) R. Br.] is typically grown. Globally, pearl millet is the sixth most important cereal grown by subsistence farmers in the arid and semi-arid regions of sub-Saharan Africa and the Indian subcontinent. Most of these agro-ecologies have low N in the root zone soil strata. Therefore, there is an immense need to identify lines that use nitrogen efficiently. A set of 380 diverse pearl millet lines consisting of a global diversity panel (345), parents of mapping populations (20), and standard checks (15) were evaluated in an alpha-lattice design with two replications, 25 blocks, a three-row plot for 11 nitrogen use efficiency (NUE) related traits across three growing seasons (Summer 2017, Rainy 2017, and Summer 2018) in an N-depleted precision field under three different N levels (0%-N₀, 50%-N₅₀, 100%-N₁₀₀ of recommended N, i.e., 100 kg ha⁻¹). Analysis of variance revealed significant genetic variation for NUE-related traits across treatments and seasons. Nitrogen in limited condition (N₀) resulted in a 27.6 and 17.6% reduction in grain yield (GY) and dry stover yield (DSY) compared to N₅₀. Higher reduction in GY and DSY traits by 24.6 and 23.6% were observed under N₀ compared to N₁₀₀. Among the assessed traits, GY exhibited significant positive correlations with nitrogen utilization efficiency (NUE) and nitrogen harvest index (NHI). This indicated the pivotal role of N remobilization to the grain in enhancing yield levels. Top 25 N-insensitive (NIS-top grain yielders) and N-sensitive (NS-poor grain yielders) genotypes were identified under low N conditions. Out of 25 NIS lines, nine genotypes (IP 10820, IP 17720, ICMB 01222-P1, IP 10379, ICMB 89111-P2, IP 8069, ICMB 90111-P2, ICMV IS89305, and ICMV 221) were common with the top 25 lines for N₁₀₀ level showing the genotype plasticity toward varying N levels. Low N tolerant genotypes identified from the current investigation may help in the identification of genomic regions responsible for NUE and its deployment in pearl millet breeding programs through marker-assisted selection (MAS).

Keywords: pearl millet, nitrogen use efficiency, genotypic variations, phenotyping, nitrogen-responsive

INTRODUCTION

Global nitrogen (N) demand, one of the most expensive farm inputs currently stands at about 117 million metric tons with a projected annual increase of ~1.5% in the future (FAO, 2019). Indian agriculture consumes over 17 million tons of N fertilizer per year. However, plants can utilize only 30–40% of the applied N for food production and the remaining (up to 60%) is lost to the environment by leaching, de-nitrification, and runoff. Surplus nitrogen pollutes freshwater streams and air which is hazardous to the majority of living species (Hickman et al., 2014; Russo et al., 2017).

Besides, excess usage of N fertilizer not only decreases the efficiency of nutrient use but also affects the rate of economic returns per unit of chemical fertilizer applied. The effect of negative environmental and economic impacts could be reduced through better agronomic practices and also by utilizing N efficient lines with improved nitrogen use efficiency (NUE) (Raun et al., 1999). Hence, improving the NUE of crop plants could help in reducing fertilizer input, increased productivity and profitability coupled with a reduced negative impact on the environment.

Globally, pearl millet (*Pennisetum glaucum* (L.) R. Br) widely grown for food and fodder is considered as the sixth most important cereal crop in terms of area. It is one of the oldest cultivated cereal crops, originated in Africa but later spread to many countries (D'Andrea and Casey, 2002; Manning et al., 2011). This millet crop is vital for food and nutritional security for the world's poorest people living in different agro-ecological zones. Besides, pearl millet has a rich nutritional profile, wide genetic diversity, high photosynthetic rates being a C₄ and also provides a healthy balanced diet, thus contributes to the economic security of poor farmers (Srivastava et al., 2019). In the Indian agriculture scenario, pearl millet is grown by poor and marginal farmers under low fertile and rainfall dependent areas often facing drought during different stages of crop growth. From the past few years, pearl millet demand has been increased for its nutritional characteristics and its adaptability to a wide range of climatic conditions (Tako et al., 2015). Since then farmers started applying a high dose of N fertilizer for maximizing grain yield (GY), but excessive usage of N fertilizer leads to low NUE of the crop along with various environmental hazards. In some parts of India, poor farmers lack knowledge about the ideal dose of fertilizer to harvest the real yield potential. Hence, yield improvement of pearl millet under low nitrogen input is indeed beneficial for economic and environmentally sustainable cultivation.

Nitrogen use efficiency is a complex trait, which is associated with various morphological, physiological, molecular, and biochemical changes in plants throughout the life cycle. For a clear understanding of this complex nature, studies of various physiological traits and their close correlation with one or more economically important traits like GY is foremost critical. Nevertheless, it will help in selecting low N tolerant/high yielding lines at different N conditions (Monostori et al., 2016). In general, plant function is always associated with chlorophyll content, which directly indicates the N status of the leaf

(Yang et al., 2014). Leaf chlorophyll content and photosynthetic capacity are appropriate benchmarks for identifying high NUE (HNUE) genotypes under low N conditions during field trials (Vijayalakshmi et al., 2015; Kiran et al., 2016). Leaf N status was usually measured by using a hand-held optical chlorophyll meter to monitor the leaf nitrogen status and chlorophyll content. Moreover, leaf area (LA) is also one of the important physiological traits, plays a critical role in enhancing plant biomass (Gianquinto et al., 2004; Monostori et al., 2016).

Nitrogen use efficiency is defined as the ability to produce GY per unit N available in the soil. NUE mainly depends on the results of two main processes, such as N uptake efficiency (NupE), and nitrogen utilization efficiency (NutE) (Good et al., 2004; Hakeem et al., 2012; Vijayalakshmi et al., 2013). NupE is the ability of the plant to take up N from the soil and NutE is the ability to use N to produce gain yield (Ladha et al., 1998; Hirel et al., 2007). Studies on genotypic variations under low and recommended N for NUE traits at both seedling and maturity stage of plants under controlled and field conditions in various crops resulted in the identification of high NUE genotypes possessing high yield sustainability under low N condition (Muchow, 1998; Inthapanya et al., 2000; Le Gouis et al., 2000; Presterl et al., 2003; Anbessa et al., 2009; Namai et al., 2009; Vijayalakshmi et al., 2015). The high NUE/N-insensitive genotypes (NIS-top grain yielders) are defined as genotypes that give more or equal GY with minimal application of N fertilizer when compared to recommended or standard N fertilizer conditions (Hawkesford, 2017). The understanding of GY and its associated NUE traits performance is still largely lagging in the pearl millet diversity panel. However, few studies had explored genotypic variations for NUE at different N levels. In one such report, 20 diverse pearl millet genotypes and few high-yielding hybrids were screened under field conditions at two N levels (Alagarswamy and Bidinger, 1987). In another study, pearl millet hybrids were grown over 3 years in western and eastern Nebraska at four different N levels to determine the optimum N rate for cultivation and concluded that the NutE trait was less responsive than the N uptake with increased N levels (Maman et al., 2006).

Nitrogen use efficiency, which is dependent on soil or external N supply, is an output of available N uptake, its efficient utilization and remobilization to grain at end of the season. Several studies revealed that improved NutE might lead to enhanced NUE under low N conditions, particularly in cereals (Moll et al., 1982; Good et al., 2004). NutE is defined as the genotype ability to assimilate and remobilize N which ultimately convert into GY. It is an essential physiological parameter that unveiled the positive relationship with GY. Whereas in pearl millet comprehensive studies are required to pinpoint the critical factors underlying NutE under different N levels. GY is a complex trait controlled by a network of multiple traits and their associations. The GY in pearl millet is a result of many yield components, such as grain number, grain weight, tiller number, and panicle number (Rai et al., 2012; Basava et al., 2019). Hence, uncovering the genetic basis of GY, and other related NUE traits under low and high N conditions is a prerequisite to understand the mechanism and also to identify

ideal NUE associated traits for selection. Thus, the current study was aimed to evaluate a diverse set of pearl millet lines for their response to GY and NUE related traits under different N levels and also to determine the suitable NUE traits for selection. This study provides useful information toward uncovering the physiological and genetic basis of GY and its related traits under low N which further facilitates the development of low N stress-resilient pearl millet cultivars.

MATERIALS AND METHODS

Experimental Details

The experiment was conducted at an N-depleted precision field of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana, India. The farm is geographically situated at an altitude of 545 m above mean sea level on 17.53°' N latitude and 78.27°' E longitude. The climate of the location is semi-arid with an average rainfall of 898 mm. The minimum and maximum temperatures ranged from 38 to 42°C which was observed during the experimental seasons. A set of 380 diverse pearl millet lines consisting of a world diversity panel, Pearl Millet Inbred Germplasm Association Panel (PMiGAP) (Sehgal et al., 2015), mapping population parents and checks were screened during three seasons (summer 2017, rainy 2017, and summer 2018) in the same field with minimal fertility and moisture gradient. Across all the three seasons, field experiments were carried out in a split-plot alpha-lattice design with 2 m and three-row plots with two replications under three N treatments [N_0 (without additional N application), N_{50} (@50 kg ha⁻¹), and N_{100} (@100 kg ha⁻¹)]. Here N source used was urea (46.4% N). Field area with three different N treatments was considered as the main plot and it was divided into six blocks (two replications for each treatment). Each N level block was divided into 25 sub-blocks, and each sub-block consisted of 16 genotypes. Nitrogen depleted plot was developed and maintained at ICRISAT over past many seasons before the inception of the experiment by withholding N application. Before initiating the experiment, important physical and chemical properties of the soil were measured. Soil samples were collected at 15 cm depth from different places and four corners of the plot; mixed composite was used to determine the soil properties (Supplementary Table 1).

Crop Management

Crop management including land preparation, sowing/planting, management of water, nutrient, disease, and pest control were taken care of throughout the crop period. Seeds were sown on the raised furrows by using a four cone planter by maintaining sufficient spacing of 60 cm from row to row and 15 cm from plant to plant. The crop was fertilized by manual broadcasting as per the treatment. N fertilization was done as per the treatments using urea (46.5%) in two equal splits at 20–25 and 40–45 days after sowing. Besides, other nutrients such as single super phosphate @250 kg ha⁻¹ and muriate of potash (60% K₂O) @50 kg ha⁻¹ was also applied

at the time of sowing. All the pearl millet genotypes were evaluated for a total of 11 traits across three consecutive growing seasons under different N regimes. The measured phenotypic traits were (1) SPAD chlorophyll content (soil plant analytical device), (2) leaf number (LN), (3) LA, (4) GY, (5) dry stover yield (DSY), (6) N percent in dry stover (NPS), (7) N uptake in grain (NUpG), (8) N uptake in stover (NUpS), (9) total nitrogen uptake (TNUp), (10) NtE, and (11) nitrogen harvest index (NHI). The phenotypic data were averaged across three seasons in each genotype for the first five traits SPAD chlorophyll content, LN, LA, GY, and DSY. For the remaining six NUE related traits data was presented as a pooled mean of two summer seasons, due to lack of quality data in the rainy season.

Physiological Traits

SPAD Chlorophyll Content (Soil Plant Analytical Device)

Soil plant analytical device chlorophyll content was measured by using Minolta Corporation's Chlorophyll SPAD-502 plus, United States. Plants from all the treatments were marked a day before with different color ribbons. Fully expanded uppermost leaves at 45 days' stage (from the day of leaf emergence) were selected and an average of three reads was recorded from a total of three plants from the middle row.

Leaf Area and Leaf Number

In order to examine the LA and LN, leaf samples were collected approximately 80 days after emergence for all the treatments and replications. Healthy plant per plot was harvested from the middle row in the field in early morning and on the same day, specific LA and number were measured by passing in all the collected leaves of the plant through LA meter LI-3100C (LI-COR, United States).

Grain and Stover Yield

At the crop harvesting stage, GY was measured after threshing sundried panicles to remove the grains from their central rachis. Harvested plants were selected from the middle row and expressed in grams (g) per 2-meter area. Similarly, DSY was also measured after drying the stover samples of all the middle row plants without panicles and expressed in grams per 2-meter area.

Nitrogen Estimation and Nitrogen Indices

Nitrogen Estimation

Nitrogen content in grain and stover samples was estimated by using the sulfuric acid-selenium digestion method. Grain and stover were made into fine powder by using clone mixture (*cyclone sample mill*) and 0.250 gm was used for nitrogen estimation in the Charles Renard Analytical Laboratory at ICRISAT, Patancheru. Pre-weighed samples were digested with the sulfuric acid-selenium and then analyzed (Sahrawat et al., 2002) by using an auto-analyzer (Skalar SAN System, AA Breda, Netherlands). Nitrogen concentration is expressed as Nitrogen% (N%).

Different Nitrogen Use Efficiencies

Total N uptake was calculated from the sum of grain uptake and stover N uptake values. Then different N efficiencies were calculated as per the formulas given by Fageria et al. (2010).

Nitrogen utilization efficiency (NUE)

$NUE \text{ (kg grain kg}^{-1} \text{ N)} = GY/TNUp$.

Nitrogen harvest index (NHI)

$NHI \text{ (\%)} = (NUPG/TNUp) \times 100$.

Statistical Analysis

The phenotypic mean values over the replications of each genotype for all the traits were prepared and the best linear unbiased estimates (BLUE) for genotypes were estimated at each and across N levels using Genstat software, 20th edition. The minimum, maximum, and BLUE of each trait at each N level were calculated. ANOVA was performed for the individual and pooled seasons by modeling individual season's residuals variance using Genstat software, 20th edition (VSN International, 2019). Fisher's *t*-test was used to ascertain the significant difference among the genotypes, treatments and interactions. Pearson's correlations coefficient analysis was performed by using pooled data of two summer seasons for all the 11 traits to identify the relationship between traits at each level of N treatments.

Cluster Analysis

Hierarchical cluster analysis was carried out using the Euclidean distance metric and UPGMA method (un-weighted paired group method and arithmetic averages)¹. By using the pooled data of three seasons treatment wise (N₀, N₅₀, and N₁₀₀) cluster analysis was performed for the five traits viz., SPAD chlorophyll content, LN, LA, GY, and DSY.

RESULTS

According to the fertilizer map, the available nitrogen (N) content of the soil sample was divided in to low, medium, and high. Therefore, low N conditions were observed when soil nitrogen content is below 280 kg/ha (critical limits of the soil maintained as per ICRISAT guide limits). The selected experimental field was not used for any cultivation from the past 5–6 years and before starting the trial, soil samples were collected across the field and undergone soil nutrient analysis. Nutrient analysis (Supplementary Table 1) revealed 0.5% organic carbon (OC), indicating a low available N in the soil and the basal N availability was common across all the N regimes before the initiation of the experiment. In addition, the soil micronutrient availability was in the medium range. The basal N availability in the soil is common for all the treatments but coming to the N₅₀ and N₁₀₀ regimes, an additional N source in the form of urea was applied as per the protocol mentioned in the methodology section. Overall, the N₀ regime was maintained strictly without any additional supply of nitrogen.

¹<http://darwin.cirad.fr/darwin>

A total of 380 pearl millet genotypes including the diversity panel, parents of mapping populations, and checks were screened for 11 physio-agronomic and NUE traits at three nitrogen levels (0, 50, and 100% of the recommended N doses) with two replications. Data recorded were averaged across the three seasons for each trait in each N level. Results revealed the native variation across the genotypes toward N response which had given the scope to identify nitrogen use efficient lines under low and high N conditions for the marginal and favorable ecologies, respectively. ANOVA results revealed significant variations among the genotypes, treatments and their interactions (season \times treatment (N levels), season \times genotype, treatment \times genotypes, and season \times treatment \times genotype). The pooled means over three seasons in each genotype in each N level for five physio agronomic traits were provided in Supplementary Table 2. The pooled mean of over two seasons for six N related traits in each genotypes were provided in Supplementary Table 3.

Physiological Traits

Key physiological traits viz., SPAD chlorophyll content, LN, and LA were recorded for all three seasons. Significant variations across the genotypes under different N regimes were noticed. Under N₀, SPAD chlorophyll content was ranged from 28.18 to 49.11 with an average of 37.85. Whereas in N₅₀, it was ranged from 32.16 to 50.1 with an average of 40.59. In N₁₀₀, the values were ranged from 29.04 to 49.07 with an average of 41.77. In N₀, SPAD chlorophyll content values were reduced by 6.8 and 9.4% compared to N₅₀ and N₁₀₀ conditions, respectively. LN was decreased under N₀ conditions by 19.3 to 29.6% compared to N₅₀ and N₁₀₀ conditions, respectively. The range of LA was lower under N₀ compared to N₁₀₀ conditions. The mean of LA was reduced significantly under the N₀ condition compared to the N₁₀₀ condition. In addition, the analysis of variance (ANOVA) results revealed that the effect of genotype, treatments and their interactions were significant ($P < 0.001$) for all the three physiological traits studied. The data was averaged across the three seasons for each trait in each N level. Descriptive statistics over the three seasons, and ANOVA results for all the physiological traits were provided in Table 1.

Agronomic Traits

Grain Yield

In N₀, the average GY of the tested genotypes was 67.04 g and ranged from 9.64 g (IP13363) to 138.48 g (IP18621). Similarly, significant genotypic variations were observed in other treatments. In N₅₀, GY values varied from 20.72 g (Tift238 D1) to 168.85 g (IP16096) with a mean of 92.61 g. Whereas in N₁₀₀, GY ranged from 20.48 g (IP13363) to 164.87 g (ICMV-IS89305) with a mean of 88.69 g. Nitrogen in limited condition (N₀) resulted in a 27.6 and 24.6% reduction in GY compared to N₅₀, and N₁₀₀, respectively. The ANOVA results indicated that the effect of genotype, treatments were significant ($P < 0.001$) and their interactions were also significant between seasons \times treatment ($P < 0.001$), seasons \times genotypes ($P < 0.001$), treatment \times genotype ($P < 0.001$), and season \times treatments \times genotype.

TABLE 1 | Summary statistics and ANOVA results from three seasons pooled data for physiological and agronomic traits under three N levels.

Traits		SPAD chlorophyll content	LN	LA	GY (g/2m ²)	DSY g/2m ²)
Range	N ₀	37.85	17.89	722.59	67.04	202.06
	N ₅₀	40.59	22.18	1212.4	92.61	246.16
	N ₁₀₀	41.77	25.40	1486.04	88.89	264.41
Mean	N ₀	28.18–49.11	8.47–34.81	202.64–2220	9.64–159.71	58.47–806.09
	N ₅₀	32.16–50.1	13.25–42.77	652.35–2967.8	20.72–236.43	62.42–0.740.78
	N ₁₀₀	29.04–49.07	10.30–41.73	363.65–2742.06	20.48–242.12	73.92–678.87
Percent reduction	N ₀ compared with N ₅₀	6.8	19.3	40.4	27.6	17.9
	N ₀ compared with N ₁₀₀	9.4	29.6	51.4	24.6	23.6
	N ₅₀ compared with N ₁₀₀	2.8	12.7	18.4	-4.2	6.9
F statistic	S	12.45***	345.29***	742.27***	580.24***	609.32***
	T	2.27 ^{ns}	40.48***	227.09***	213.16***	55.08***
	S × T	0.67 ^{ns}	15.21***	67.86***	67.77***	93.49***
	G	5.94***	7.44***	8.48***	9.48***	14.97***
	S × G	4.65***	5.19***	7.77***	6.87***	7.54***
	T × G	2.49***	3.5***	3.6***	2.86***	4.01***
	S × T × G	2.17***	2.87***	4.3***	2.32***	2.73***

SPAD, Chlorophyll content; LN, leaf number; LA, Leaf area; GY, grain yield; DSY, dry stover yield; G, Genotype; S, Season; T, treatment.

F statistic values with asterisk indicate significant at levels of 0.001 (***) and NS denote for non-significant.

Based on GY data, top 25 (N-insensitive (NIS-Top grain yielders) and least 25 (N-sensitive (NS-Poor grain yielders) genotypes were identified under N₀ conditions (Table 2). Out of 25 NIS lines, nine genotypes (IP10820, IP17720, ICMB01222-P1, IP10379, ICMB89111-P2, IP8069, ICMB90111-P2, ICMV-IS89305, and ICMV221) were common in the top 25 lines at the N₁₀₀ level which shows the genotype plasticity toward N₀ and N₁₀₀ conditions (Figure 1). Similarly, the top 25 high yielding genotypes were identified in N₁₀₀. The 25 least grain yielding genotypes [N-sensitive (NS)] in N₀ and N₁₀₀ conditions were presented in Figure 2.

Furthermore, agronomic data was provided for the top five lines viz., IP 18621, IP 10820, IP 17720, IPC 804, and ICMB01222-P1, out of selected 25 NIS lines (Supplementary Table 4). Interestingly, days to 50% flowering (DF50%) and plant height (PH) was more in all the tested genotypes at N₀, indicating the genotype efficiency in terms of agronomic traits. These genotypes can be used as parents in the pearl millet breeding program to develop N efficient genotypes under low N input conditions.

Dry Stover Yield

In N₀, DSY ranged from 58.47 g (LGD-1-B-10) to 806.19 g (IP8863) with a mean of 256.3 g. Whereas in N₅₀, values ranged from 62.42 g (IP13363) to 740.78 g (IP11584) with a mean of 246.16 g. In N₁₀₀, the average DW of 380 genotypes was 264.41 g, and values varied from 73.92 g (Tift23D2B1-P1-P2) to 678.87 g (IP8786). Application of N significantly increased DSY by 12.6% in N₅₀ and 18.02% in N₁₀₀ compared to the N₀ condition. Significant differences were observed for dry stover weight among the genotypes ($P < 0.001$), and interactions were also significant between seasons × treatment ($P < 0.00$), seasons × genotypes ($P < 0.001$), treatment × genotype ($P < 0.001$), and season × treatments × genotype.

Interestingly, out of the top 25 high dry stover yielding lines, nine genotypes (IP14398, IP10579, IP15857, IP10394, IP17125,

IP13608, IP6098, IP12020, and IP11584) were common in the top 25 lines at N₁₀₀ condition (Figure 3). At both N levels, the top 25 poor-performing dry stover genotypes were also identified (Figure 4).

TABLE 2 | Top 25 NIS and NS lines under low nitrogen (N₀) with contrasting NUE grain yield derived from the pooled mean over three seasons.

S. No	Genotypes	NIS	Genotypes	NS
1	IP18621	138.42	IP13363	9.64
2	IP10820	131.19	IP3110	11.54
3	IP17720	130.81	IP9282	13.56
4	IPC804	127.12	IP4378	15.27
5	ICMB01222-P1	119.13	IP18500	15.43
6	IP10379	116.92	IP15344	15.75
7	IP3108	115.54	IP14398	18.45
8	IP11577	115.52	IP14311	18.72
9	IP16096	114.84	IP7941	19.73
10	IP5560	113.03	IP5031	20.88
11	ICMB89111-P2	110.84	IP10456	21.15
12	IP8069	109.29	IP13608	21.84
13	IP3865	109.15	AIMP92901-S1-15-1-2-B-P03	22.09
14	ICMV-IS92222	108.9	IP6310	22.31
15	ICMB90111-P2	108.34	IP4979	24.2
16	IP6099	107.22	IP6869	24.92
17	IP13016	104.61	ICMB90111-P5	26.13
18	IP17028	103.4	IP19584	26.53
19	Tift238D1-P158	103.31	IP7930	27.56
20	IP16403	102.3	IP11593	29.33
21	ICMV-IS89305	101.87	RIB334/74-P1	29.81
22	IP6109	101.32	IP15070	29.94
23	IP6060	100.37	IP5438	30.23
24	IP22424	100.3	IP8000	31.19
25	ICMV221 = ICMV88904	100.22	Tift186	31.22

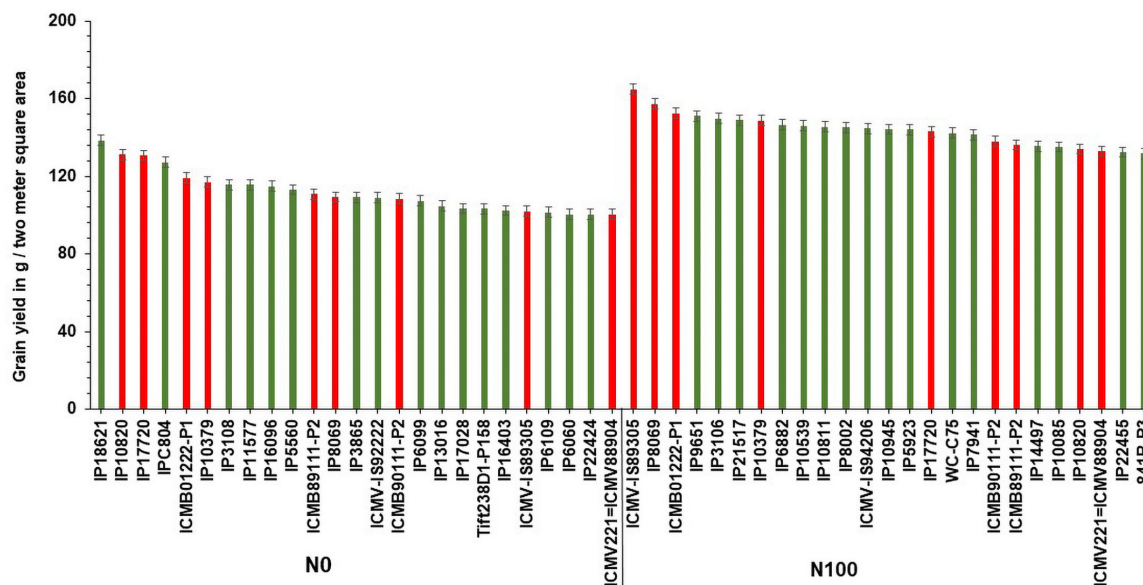


FIGURE 1 | Top 25 high grain yielding genotypes under N₀ and N₁₀₀ conditions. Total 9 genotypes (in red color) viz. IP 10820, IP 17720, ICMB 01222-P1, IP 10379, ICMB 89111-P2, IP 8069, ICMB 90111-P2, ICMV IS89305, and ICMV 221 exhibited high grain yield in both N₀ and N₁₀₀ conditions.

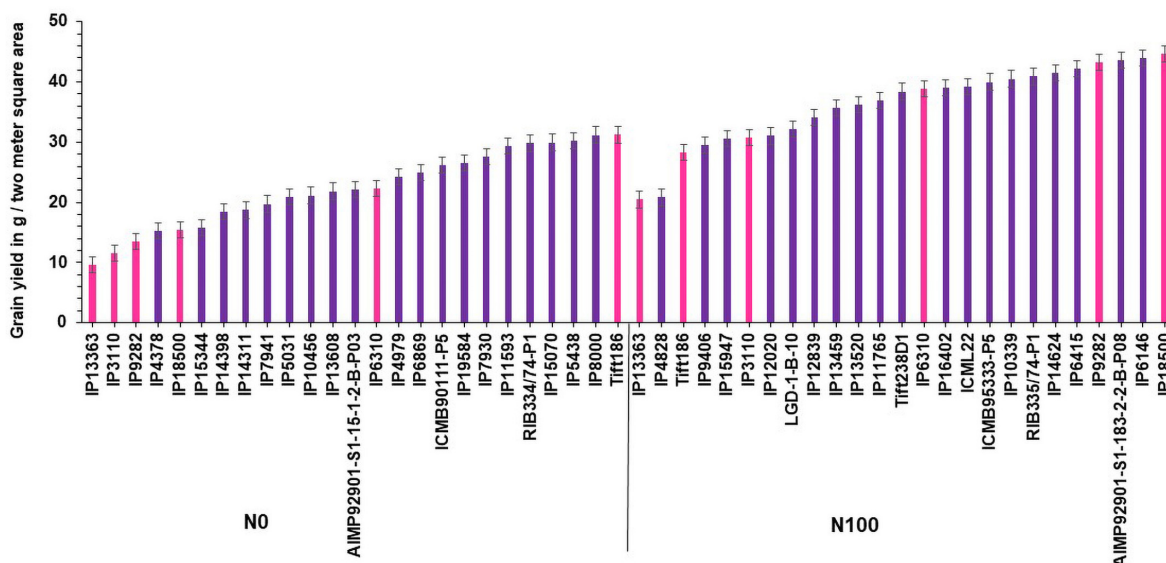


FIGURE 2 | Top 25 poor yielding genotypes for grain yield under N₀ and N₁₀₀ conditions.

In addition, three high yielding dual purpose genotypes (IP5560, IP15857, and IP17125) were identified in the top 25 GY and DSY genotypes under N₀. Whereas in N₁₀₀, only one common high yielding genotype (IP3106) was identified. Descriptive statistics over the three seasons, and ANOVA results for the agronomic traits were provided in Table 1.

N Content in Grain and Stover

Nitrogen percent in grain (NPG) varied from 1.18% (IC804) to 2.43% (Tift 186), with an average of 1.61% in N₀ condition.

Whereas in N₅₀, NPG ranged from 1.19% (9444) to 2.43% (IP4942) with a mean of 1.67%. Under N limiting conditions NPG was reduced by 3.6 and 11.4% as compared with N₅₀ and N₁₀₀ conditions, respectively. In N₅₀, an average mean of NPG decreased by 8.24% as compared to N₁₀₀. The response of genotypes in the N₁₀₀ condition varied significantly from 1.27% (IP10539) to 2.59% (IP4828) with a mean of 1.82%. Significant differences were observed in NPG among the genotypes ($P < 0.001$), treatments ($P < 0.001$), and interactions and are also significant

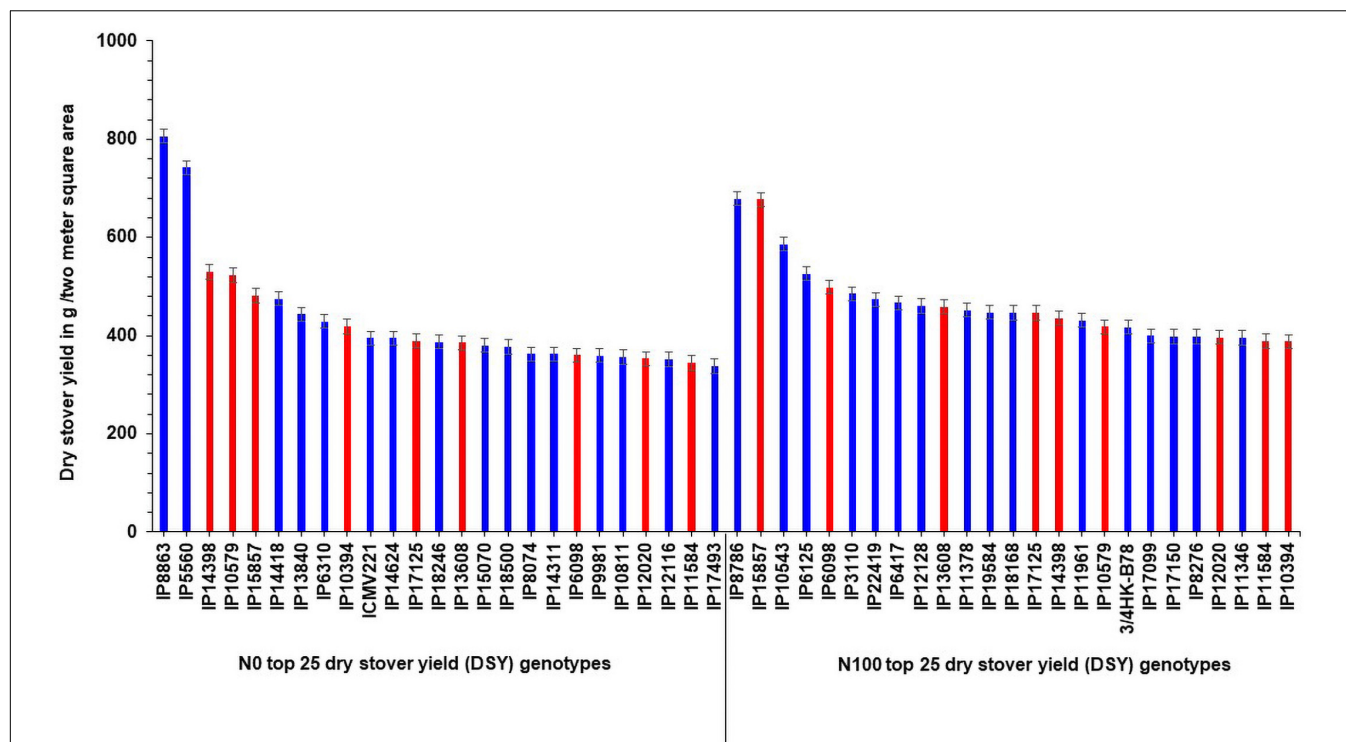


FIGURE 3 | Top 25 high dry stover yielding (DSY) genotypes under N_0 and N_{100} conditions. Total 9 genotypes (highlighted in red color) viz., IP 14398, IP 10579, IP 15857, IP 10394, IP 17125, IP 13608, IP 6098, IP 12020, and IP 11584 exhibited high dry stover yield (DSY) in both N_0 and N_{100} conditions.

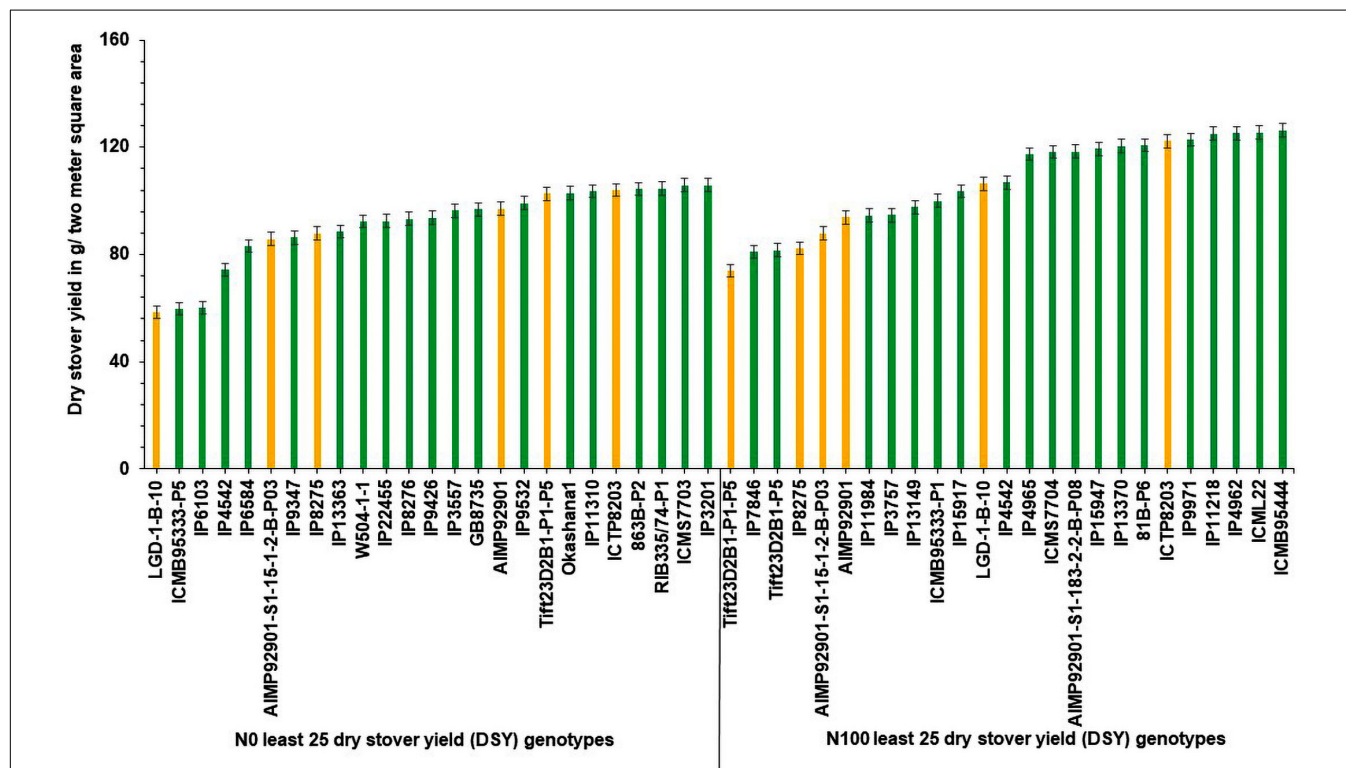


FIGURE 4 | Least performing genotypes for dry stover yield (DSY) under N_0 and N_{100} conditions.

between seasons \times treatment ($P < 0.00$), seasons \times genotypes ($P < 0.001$), and season \times treatments \times genotype except for treatment \times genotype (0.677).

Genotypic variations found in NPS in N_0 and the minimum and maximum value of NPS ranged from 0.42% (IP8786) to 1.08% (ICMB90111-P5) with a mean of 0.61%. The average NPS was more in N_0 compared to N_{50} and N_{100} conditions. Significant variations were observed in all the interactions except season \times treatment (0.6771) and treatment \times genotypes (0.0102).

Nitrogen Uptake for Dry Stover

Genotypic variations were observed for NUpS across the N levels. The mean of all genotypes was 1.65 kg ha^{-1} ranged from 3.34 kg ha^{-1} (ICMB9533-P5) to 8.03 kg ha^{-1} (IP5560) and decreased by 9.8 and 14.45% in N_0 compared to N_{50} and N_{100} , respectively. Whereas, in N_{50} , the values varied from 0.51 kg ha^{-1} (AIMP92901-S1-15-1-B-P3) to 8.03 kg ha^{-1} (IP10543) with a mean of 1.83 kg ha^{-1} and decreased by 5.67% when compared with N_{100} . The minimum and maximum values of NUpS varied from 0.38 kg ha^{-1} (IP155363) to 1.08 kg ha^{-1} (IP11584) with a mean of 6.59 kg ha^{-1} under N_{100} condition.

Total Nitrogen Uptake

Application of N fertilizer resulted in an increase of TNUP and also revealed significant genotypic variations among the genotypes with in the treatments. Under N_0 , the values ranged from 1.09 to 10.47 with an average of 3.03 kg ha^{-1} and decreased by 14.4 and 21% as compared with N_{50} and N_{100} , respectively. Whereas in N_{50} and N_{100} , values varied from 1.08 (IP11229) to

9.57 (843B) and 0.58 (IP13363) to 8.97 (IP11584) with a mean of 3.54 and 3.83 kg ha^{-1} , respectively. The results of ANOVA for TNUP indicated significant effects for all the interactions except treatments.

Nitrogen Utilization Efficiency

Large variation was observed for NUtE of genotypes, ranged from 2.40 (IP3110) to 50.60 (IP3557) with an average of 30.81 under N_0 and the efficiency was reduced as compared with N_{50} and N_{100} by 7.3 and 5.95%, respectively. Whereas in N_{50} , values varied from 10.63 to 50.86 with a mean of 33.22, and a 7.57% reduction was observed as compared with the N_{100} condition. The application of N fertilizer (N_{100}) resulted in a significant increase in NUtE. The minimum and maximum NUtE were observed in IP31110 (10.06) and IP10456 (56.25), respectively with an average of 32.76. Significant differences were observed among the genotypes ($P < 0.001$), treatments ($P < 0.001$), and seasons \times treatments ($P < 0.001$), seasons \times genotypes ($P < 0.001$), treatment \times genotypes and interactions were also significant between season \times treatments \times genotype.

Nitrogen utilization efficiency under N_0 was compared with NUtE under N_{100} to determine the genotypic efficiency and their responsiveness to N. Based on the NUtE, the genotypes were classified into four groups (Figure 5) viz., (1) N efficient non-responsive (NENR), (2) N efficient responsive (NER), (3) N responsive inefficient, and (4) N inefficient non-responsive (Rengel and Graham, 1995; Worku et al., 2007). In this study, NUtE data recorded in 374 genotypes across different N levels. The average NUtE of 374 genotypes under N_0 (30.81) and N_{100}

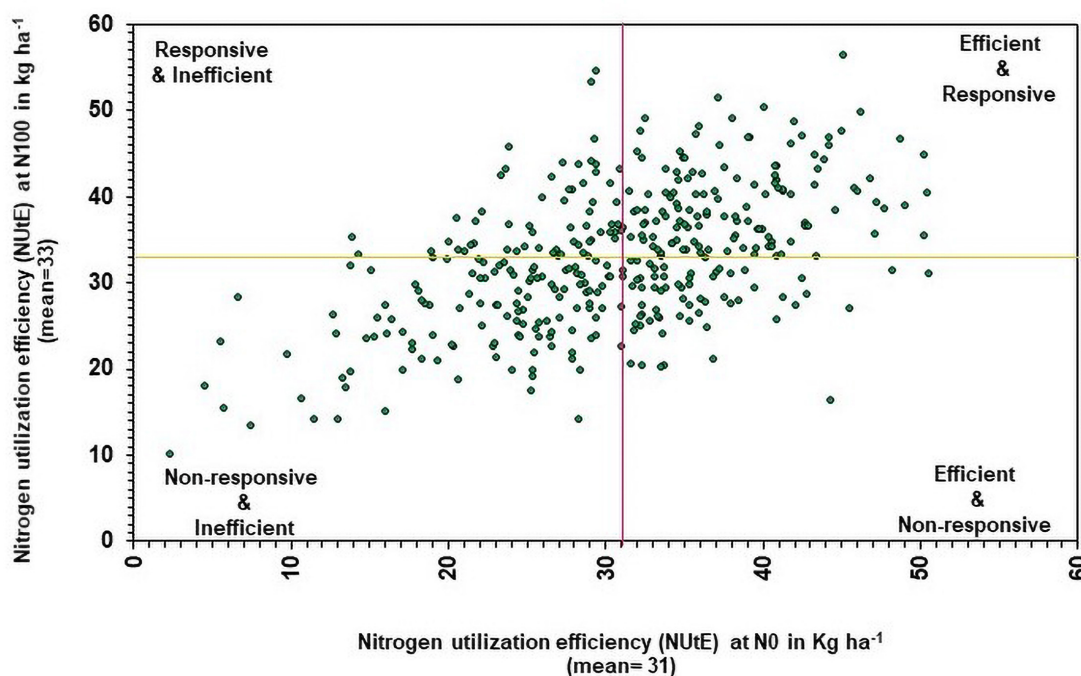


FIGURE 5 | Relationship between genotypes performance of nitrogen utilization efficiency (NUtE) under low (N_0) and recommended N (N_{100}) conditions. The yellow line represents the mean of NUtE at N_{100} and the purple line represents the mean of NUtE at N_0 (pooled data from two summer seasons 2017 and 2018).

TABLE 3 | ANOVA for N-related traits under three nitrogen levels during Summer 2017 and 2018.

Traits		NPG (%)	NPS (%)	NUpS (kg ha ⁻¹)	TNUp (kg ha ⁻¹)	NUtE (kg ha ⁻¹)	NHI (%)
Range	N ₀	1.18–2.43	0.42–1.08	0.34–8.03	1.09–10.47	2.42–50.6	5.46–74.52
	N ₅₀	1.19–2.43	0.4–0.94	0.51–6.27	1.08–9.57	10.63–56.86	21.07–84.02
	N ₁₀₀	1.27–2.59	0.32–1.03	0.38–6.59	0.58–8.97	10.06–56.27	23.51–78.89
Mean	N ₀	1.61	0.61	1.65	3.03	30.81	47.56
	N ₅₀	1.67	0.56	1.83	3.54	33.22	53.00
	N ₁₀₀	1.82	0.56	1.94	3.83	32.76	56.24
Percent reduction	N ₀ compared with N ₅₀	3.6	–8.9	9.8	14.4	7.3	10.3
	N ₀ compared with N ₁₀₀	11.54	–8.93	14.95	20.89	5.95	15.43
	N ₅₀ compared with N ₁₀₀	8.24	0.00	5.67	7.57	–1.40	5.76
F statistic	S	1808.86***	717.71***	7105.54***	12241.28***	129.2***	45.79
	T	39.22***	12.04***	1.75 ^{ns}	2.52 ^{ns}	43.13***	43.8***
	S × T	0.39 ^{ns}	5.39**	148.57***	212.2***	172.45***	16.36***
	G	9.34***	3.48***	7.7***	6.66***	9.86***	8.33***
	S × G	5.19***	2.67***	6.44***	4.58***	4.5***	4.72***
	T × G	1.85***	1.15**	3.91***	3.07***	2.09***	2.47***
	S × T × G	1.47***	1.1 ^{ns}	2.26***	1.85***	2.09***	2.31***

NPG, N percent in grain; NPS, N percent in dry stover; NUpS, N uptake in stover; TNUp, total nitrogen uptake; NUtE, nitrogen utilization efficiency; NHI, nitrogen harvest index; G, Genotype; S, Season; T, treatment.

F statistic values with asterisk indicate significance at the level of 0.01 (**) and 0.001 (***) and NS denotes no significance.

condition (32.76) were considered as cut off for the identification of genotypic efficiency and responsiveness for N use. Under low N, above-average genotypes were considered as efficient and below-average genotypes were considered as inefficient. Similarly, in N₁₀₀, above-average genotypes were considered as responders, and below-average genotypes are considered as non-responders. Overall, efficient genotypes are higher in the utilization of absorbed N over inefficient genotypes.

Nitrogen Harvest Index (NHI)

The nitrogen harvest index varied from 5.46% (IP3110) to 74.52% (ICMB 95333-P5) with an average of 47.56% in N₀ condition and a reduction of about 10.3 and 15.43% compared to N₅₀ and N₁₀₀ condition, respectively. Whereas in N₅₀, NHI values ranged from 21.10% (IP10488) to 84.02% (IP5923) with a mean of 53.0%. The response of genotypes under the N₁₀₀ condition varied significantly and ranged from 23.51% (IP12020) to 78.89% (IP6110) with a mean of 56.24%. Further, the NHI reduction under N₀ was about 10.3 and 15.43% as compared with N₅₀ and N₁₀₀, respectively. In N₅₀, a reduction of NHI about 5.46% was observed compared to the N₁₀₀ condition. Significant differences were observed among the genotypes ($P < 0.001$), treatments ($P < 0.001$), and interactions were also significant between seasons × treatment ($P < 0.001$), seasons × genotypes ($P < 0.001$), and season × treatments × genotype except for the treatment × genotype (0.677) (Table 3).

Phenotypic Correlation

Correlation coefficient analysis was performed to identify the interrelation among the traits for all the N levels. The complete list of correlation coefficient values among the traits in each N level is provided in Figure 6. GY showed a significant positive correlation with TNUp, NUtE, and NHI. The results revealed

that the GY of pearl millet was increased through the selection of higher NUtE and NHI lines. Furthermore, GY negatively correlated with DSY, NPS, NPG, and NUpS at both N₀ and N₁₀₀ conditions. The results suggested that the higher GY was, the lower DSY, NPS, NPG, and NUpS. Further DSY showed a negative significant correlation with all the measured traits except NUpS and TNUp across N levels. Interestingly NUtE showed a positive correlation with GY and NHI across N levels. SPAD chlorophyll content was found to be positively correlated with all traits except LA, DSY, and NUpS across all the treatments. Further LA was negatively correlated with GY, NUtE, and NHI across the treatments.

Cluster Analysis

Hierarchical cluster analysis of 380 pearl millet genotypes revealed large genotypic variation in five traits viz., SPAD chlorophyll content, LN LA, GY, and DSY. This analysis is useful for grouping genotypes for NUE. Therefore, clustering 380 genotypes based on five traits identified two to three major clusters. Based on this, genotypes were classified into high, medium, and low-performing genotypes. At both N levels, clusters I and II contained better-performing genotypes and also found the least dissimilarity among the genotypes. Interestingly few common genotypes viz., IP 15536, IP 6098, IP11961, IP8863, IP11378, IP 20679, and IP 12138 were identified in N₀ and N₁₀₀ levels in the first two clusters, implying that genotype plasticity toward N under N₀ and N₁₀₀ conditions (Figures 7, 8). Cluster III contained more number of medium performing genotypes that form a maximum number of sub clusters at both N levels. Finally, the least performing genotypes were present in Cluster IV. Overall, the maximum dissimilarity (from all these matrices) was observed between cluster I and IV.

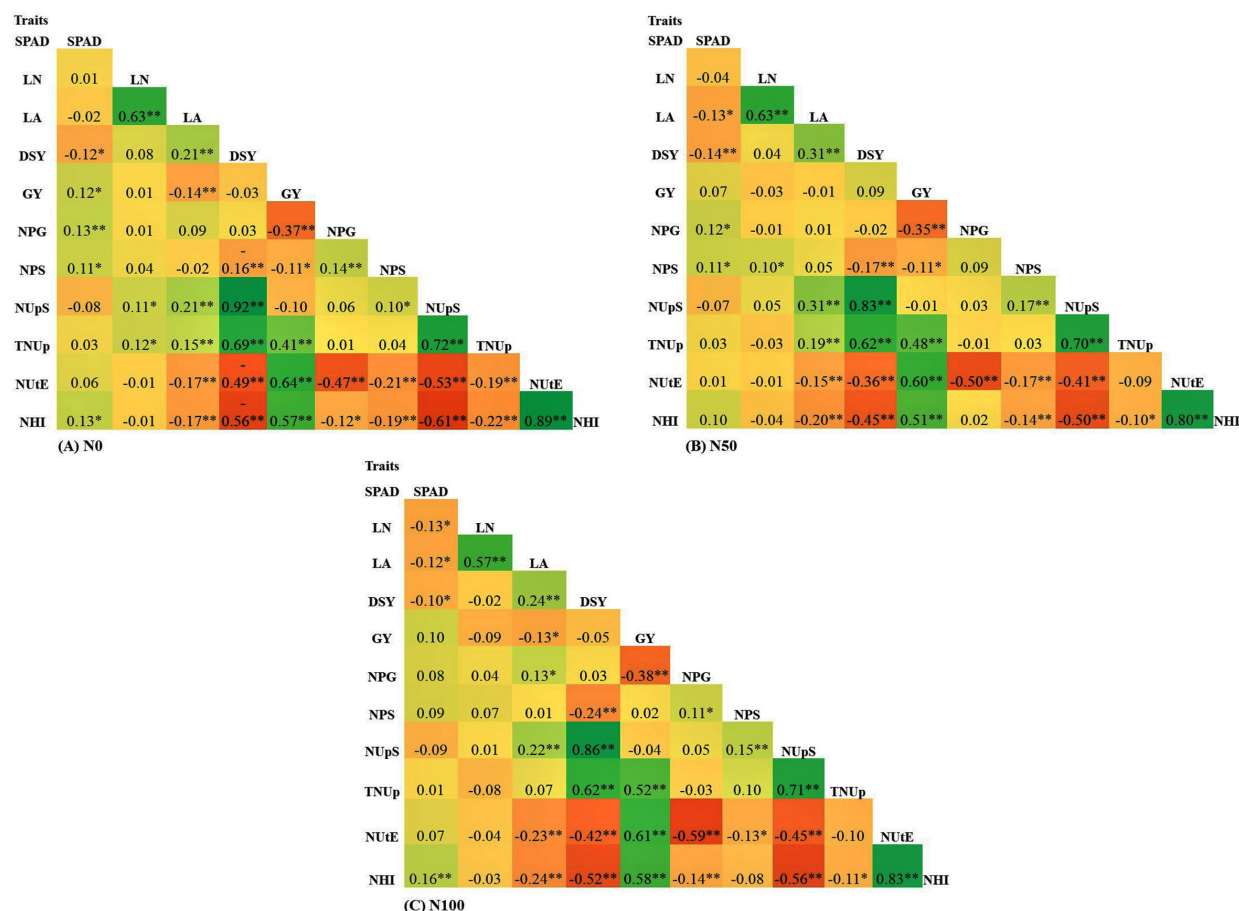


FIGURE 6 | Pearson correlation for traits underlying NUE in 380 global diverse pearl millet genotypes (PMiGAP) under three different nitrogen (N) regimes carried out in 2017/18 (data pooled from two summer seasons). **(A)** Correlations among eight traits at N₀ condition (N₀). **(B)** Correlations among eight traits at N₅₀ condition. **(C)** Correlations among eight traits at N₁₀₀ condition. * indicates < 0.05 significance, ** indicates < 0.01 significance.

DISCUSSION

Nitrogen (N) is an essential macronutrient required for plant growth and is often a limiting factor for crop yield. From the past 50 years of agriculture, augmented food production was attained by the extensive application of N fertilizer in combination with N-responsive cultivars across the globe. Excessive application of N fertilizer is becoming expensive which accounts for the loss of economic profit to the farmers along with negative impacts on the environment (Raun et al., 2002; Hawkesford and Griffiths, 2019). Hence to overcome this problem, a clear understanding of genotype behavior, identification and development of genotypes (without compromising GY) with high NUE under low N conditions is a paramount need for improving NUE.

Therefore, the present study aimed to identify the N-Sensitive (NS) and N-insensitive (NIS) genotypes and also to determine the traits regulating low N tolerance. However, accurate and reliable phenotyping under low N input is challenging and influenced by the genotype (G), environment (E), and G × E interactions (Chen et al., 2014; Rao et al., 2018). Moreover, there are limited studies available in pearl millet toward the identification

of nitrogen insensitive genotypes under low N conditions at field level. Based on the current knowledge, this study presents the first report of genetic variations for NUE in the global association panel, Pearl Millet inbred Germplasm Association Panel (PMiGAP) (PMiGAP collected from a large set of diverse germplasm and breeding lines) and from mapping population parents. Notably, very few pearl millet breeding programs are targeting the development of low N tolerance traits which are a must for sustainable agriculture with minimal negative impacts on the environment. Components of NUE were studied in pearl millet hybrids at two N levels found that the efficient genotype (Souna B) had a 32% of higher NUE value than the N inefficient Indian genotype (BJ104) (Alagaraswamy and Bidinger, 1987). Another study revealed that NUtE contributes more to genetic variation in NUE (Maman et al., 2006).

In this study, a total of 380 diverse pearl millet genotypes were characterized for SPAD chlorophyll content, LN, LA, GY, DSY, and NUE related traits at three different N levels under field conditions. Notably, diverse responses have been observed among the genotypes across N levels, despite a similar growth conditions and an equal amount of N fertilizer in a given N

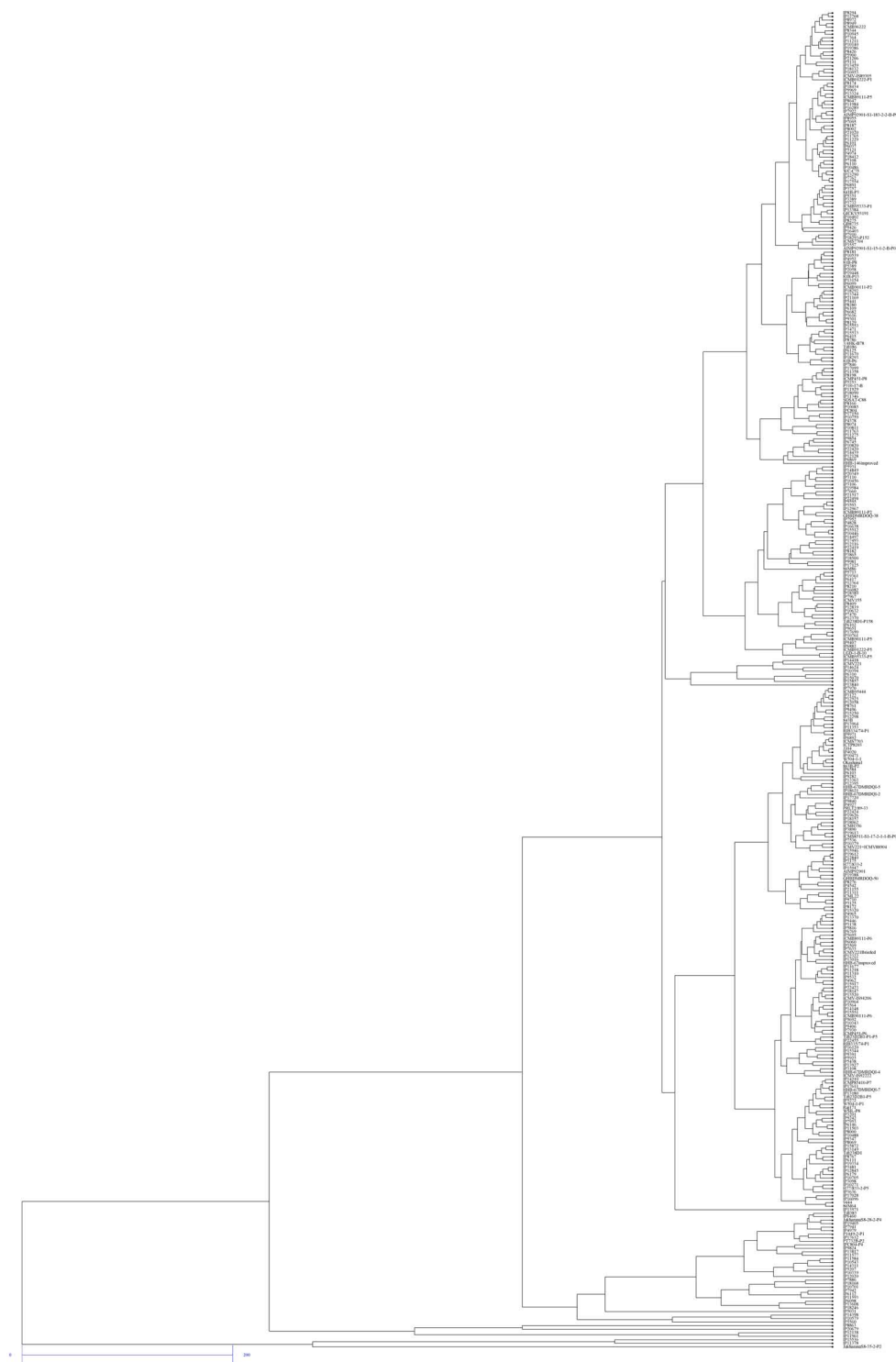


FIGURE 7 | Dendrogram of 380 pearl millet genotypes for five traits under low nitrogen conditions (N_0) by the UPGMA method.

level. These observed genotypic variations purely reveal the genotype plasticity toward traits. This set of lines used in this study which are collected from different regions across the globe and these lines were used for various abiotic stress studies

to identify the tolerant genotypes. Previous studies in rice, maize, wheat, and oilseed rape, etc., have established significant genetic variation for NUE-related traits with large germplasm panels, hybrids, open-pollinated and recombinant inbred line

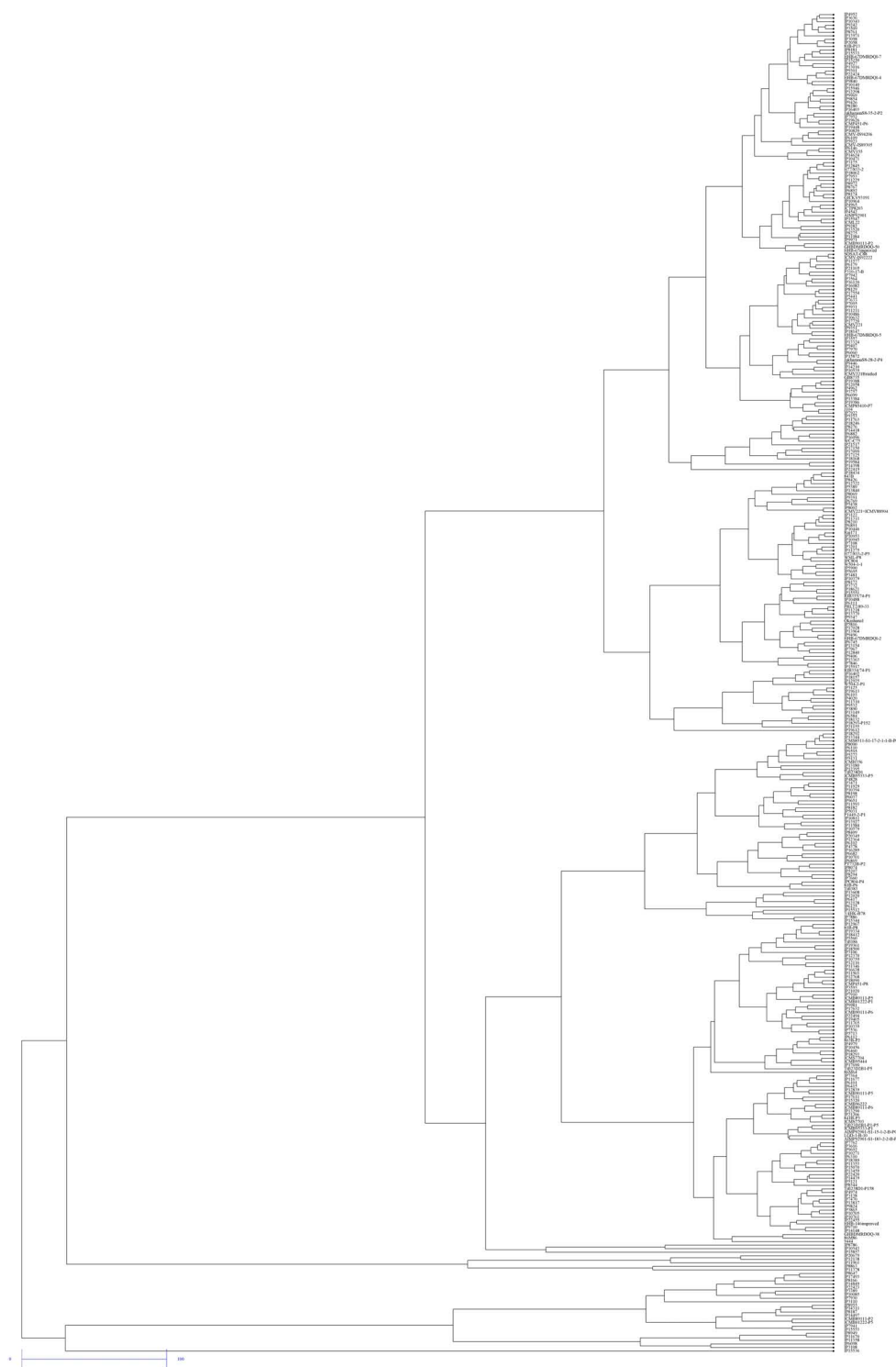


FIGURE 8 | Dendrogram of 380 pearl millet genotypes for five traits under recommended conditions of nitrogen (N_{100}) by the UPGMA method.

populations (Chen et al., 2014; Li et al., 2015; Vijayalakshmi et al., 2015; Ertiro et al., 2017; He et al., 2017; Rao et al., 2018). Generally, the results of ANOVA explained statistical differences among the treatments and genotypes and their interactions. In

the current study, ANOVA results revealed significant variations among the genotypes and treatments for all measured traits except a few traits (Tables 1, 3). These results were indicating that the environment under different N levels was a crucial factor

in explaining the genotypic variance of GY and its related traits, which are strongly influenced by the relative contribution of $G \times E$ interactions under field evaluation. The results concur with other reported field experiments in rice (Srikanth et al., 2016) and wheat (Sial et al., 2005; Belete et al., 2018).

In the present study, SPAD chlorophyll content significantly increased with nitrogen application, which might be sufficient availability of N in the leaf. Similar results of increased SPAD chlorophyll content by increasing N application was reported by Kitajima and Hogan (2003), Pramanik and Bera (2013). The relation between the slow or controlled release of N fertilizers supporting more N absorption and related physiological mechanisms has been well studied in different crops (Li et al., 2003; Long et al., 2013). LA of all the pearl millet genotypes was significantly increased with N application which might be attributed to translocation of N to leaves, which brings variation in plant architecture and leaf internal structure (Singh et al., 2004; Mhaskar et al., 2005). Reduction in LN under N_0 condition was observed, due to deprivation of N, thereby reduced source size and hindering the plant development.

Across the year on average, GY ranged from 9.64 g to 159.71 g (N_0), 20.72 g to 236.43 g (N_{50}), and 20.48 g to 242.12 g (N_{100}), indicating the genotypic variability in GY at different N regimes. In general, GY increased was correlated with the N application rate, which might be due to sufficient nitrogen availability. Recently similar kind of results was reported in a diverse set of foxtail millet genotypes (Bandyopadhyay et al., 2020). In wheat, GY and quality traits were evaluated by Šarčević et al. (2014) at low and normal N and found that GY was reduced by 10% under low N condition compared to normal condition. The current results also are in tune with the earlier reports as there was an increased GY with increase N levels. Hawkesford and Griffiths (2019) reported that split application of N is the best option to recover the maximum applied N in the form of harvested grain/increased GY and found a strong significant relationship between GY and physiological traits.

Remarkably, 25 high yielding genotypes were identified under the N_0 condition and considered as N-insensitive (NIS). Out of 25, nine genotypes (IP10820, IP17720, ICMB01222-P1, IP10379, ICMB89111-P2, IP8069, ICMB90111-P2, ICMV-IS89305, and ICMV221 = ICMV88904) are common in the top performing 25 lines under N_{100} condition, which shows the genotype plasticity toward varying N levels in these genotypes. These selected genotypes are good genetic resources to breed for tolerance to low N conditions. Top 25 high yielding genotypes in N_{100} can be considered to have the best acceptances for cultivation where soils are fertile and when followed the ideal N levels for cultivation. These identified genotypes could be used to improve GY, along with higher NUE. Numerous studies have reported utilization of N efficient lines with enhanced GY in the farmer fields which may help to reduce fertilizer input as well as increase the profitability of farm operations (Würschum, 2012; Vijayalakshmi et al., 2015; He et al., 2017). Likewise, selected NIS lines with high NUE will certainly play a role in reducing environmental pollution and could increase economic profit to farmers.

Pearl millet is an ideal fodder crop for feeding livestock and has unique features like dry season crop, production environments having low nitrogen, high photosynthetic

efficiency, and high dry forage capacity (Govintharaj et al., 2018). Several reports depict the dry fodder demand and it would require approximately 568 million tonnes across the globe by the year 2030 (Govintharaj et al., 2018). To tackle this problem, high fodder yielding genotypes need to be identified. Generally, most of the genotypes have lower dry matter production under low N conditions, hence identification of genotypes with high dry matter production under low N conditions is important for sustainable fodder production. In this framework, the top and least performing 25 genotypes for DSY were identified under N_0 and N_{100} conditions (Table 4).

These contrasting genotypes can be further exploited in the breeding program and molecular studies to develop varieties with high NUE for biomass. Irrespective of N, the common nine genotypes (IP14398, IP10579, IP15857, IP10394, IP17125, IP13608, IP6098, IP12020, and IP11584) were identified and can produce more dry matter production under both N_0 and N_{100} conditions. These genotypes need further physiological and molecular characterization under low nitrogen conditions to understand the molecular basis for low nitrogen tolerance and biomass production. Overall, observed genetic variability in pearl millet genotypes are suited well for improving DSY, can be utilized in the breeding program to enhance livestock productivity. Studies on various crops to determine the relationship between nitrogen levels and biomass traits have shown that N fertilizer application increased the biomass related traits, including DSY (Chan-Navarrete et al., 2014). The present

TABLE 4 | Dry stover yield (DSY) for top and least 25 genotypes under N_0 condition (derived from pooled data of three seasons).

S. No	Genotypes	Top 25 DSY	Genotypes	Least 25 DSY
1	IP8863	806.09	LGD-1-B-10	58.47
2	IP5560	741.82	ICMB95333-P5	59.72
3	IP14398	529.71	IP6103	60.13
4	IP10579	523.27	IP4542	74.33
5	IP15857	481.52	IP6584	83.17
6	IP14418	475.12	AIMP92901-S1-15-1-2-B-P03	85.84
7	IP13840	443.79	IP9347	86.42
8	IP6310	428.9	IP8275	87.84
9	IP10394	419.14	IP13363	88.54
10	ICMV221	395.32	W504-1-1	92.31
11	IP14624	395.1	IP22455	92.39
12	IP17125	389.29	IP8276	93.25
13	IP18246	387.36	IP9426	93.73
14	IP13608	386.07	IP3557	96.39
15	IP15070	380.06	GB8735	96.83
16	IP18500	377.18	AIMP92901	97.12
17	IP8074	362.87	IP9532	99.22
18	IP14311	362.63	Tift23D2B1-P1-P5	102.62
19	IP6098	360.08	Okashana1	102.92
20	IP9981	359.06	IP11310	103.67
21	IP10811	356.19	ICTP8203	104
22	IP12020	353.03	863B-P2	104.48
23	IP12116	351.76	RIB335/74-P1	104.52
24	IP11584	344.58	ICMS7703	105.77
25	IP17493	337.88	IP3201	105.86

results were tuned with these observations as DSY was more in N_{100} condition than in N_0 condition; the increasing quantity was up to the significant levels. Another study revealed wide genetic variations for biomass traits and stover nitrogen in pearl millet germplasm (Gupta et al., 2015). Few studies revealed that significant genetic variation present in pearl millet for biomass traits in dual-purpose hybrids and open-pollinated varieties (OPVs), populations, and top cross hybrids (Bidinger et al., 2010; Blümmel et al., 2010; Rai et al., 2012). More importantly from this study, dual-purpose genotypes were identified and these can be used as a reference set for the developing high grain and stover yielding lines.

In the present study, a total of 11 traits were recorded for all three seasons. However, due to heavy rain, we are unable to generate quality data for the rainy season. Hence to avoid misinterpretation of results/findings, data for only six NUE traits were pooled from two summer seasons and presented. Moreover, the NPG was slightly increased with the nitrogen application and genetic differences for the traits were also observed. These results are in tune with the report of Belete et al. (2018), who found the higher NPG with increased N levels. NPG in N_0 was reduced by 3.6 and 11.4% compared to N_{50} and N_{100} , respectively. Whereas, in N_{50} condition, the mean NPG decreased by 8.24% as compared to N_{100} . These results are similar to the finding of Lopez-Bellido et al. (2004), Arduini et al. (2006) who found the genetic variations among wheat genotypes for grain N concentration and increased NPG with increased N levels. In the case of the NPS, the trend was contrasting to that of NPG. Herein, NPS was slightly more in N_0 as compared with N_{50} and N_{100} but was not up to a significant level. These results concur with the earlier reports of Belete et al. (2018) who found higher NPS with increasing N levels.

Nitrogen utilization efficiency is defined as the genotype ability to assimilate and remobilize N ultimately to produce the GY (Rengel and Graham, 1995; Fageria et al., 2008). Determination of the genetic variations in NUtE is essential for the selection of efficient genotypes and can be used further in breeding programs to develop low N tolerant material. The concept of genotypes grouping is used widely in nutrient use efficiency (Fageria and Baligar, 2003). Based on NUtE efficiency data in N_0 versus N_{100} , genotypes are classified into four groups viz., N efficient non-responsive, N efficient responsive (NER), N responsive inefficient, and N inefficient non-responsive under N_0 and N_{100} condition, respectively. In N_0 NUtE data, a total of 198 efficient and 176 inefficient genotypes were identified and out of 198 efficient genotypes, 58 genotypes were falling in the first desirable group which is N efficient non-responsive. These genotypes are exhibiting a progressive performance under low N. This may enable breeders to develop efficient genotypes under low input environments for pearl millet breeding activities. The remaining 140 genotypes were falling in the next most desirable group, NER and these genotypes were exhibiting a progressive response to increased N availability. The NER genotypes identified from the present investigation could be the prospective targets for selection toward the genetic improvement of pearl millet for N utilization. Interestingly, in the identified 25 NIS genotypes, 20 genotypes which are falling under N efficient responsive group were showed progressive performance in terms of efficient and

responsive use of nitrogen. Genotypes with more NUtE could produce high GY per unit of N consumption (Moll et al., 1982). Various studies on NUtE have already reported that breeding for efficient genotypes under low N could be achievable with high NUtE (Fageria, 2014; He et al., 2017). The third group, responsive inefficient genotypes, can be used in breeding programs. The rest of the genotypes fall into the fourth group and these are less desirable from the NUE point of view. Interestingly, few common genotypes were identified from the non-responsive and inefficient genotypes. Overall, efficient genotypes are higher in the utilization of absorbed nutrients than inefficient genotypes. Vast genotypic variations in NUtE have been reported under field/pot screening in various wheat genotypes and other crops by several researchers (Bouchet et al., 2014; Ma et al., 2015).

The nitrogen harvest index indicates the level of efficiency of plants to use acquired total N for grain formation (He et al., 2017). In the present study, significant genotypic variation was observed for NHI at both N_0 and N_{100} conditions. NHI was positively correlated with NUtE and GY in all three N levels. These findings concur with previous studies in oilseed rape (Ulas et al., 2013; Stahl et al., 2016) and wheat (Monostori et al., 2017).

In the current study, the Pearson correlation coefficient analysis for the GY and NUE related traits indicated that NUtE trait has the largest contribution to the GY followed by NHI and TNUp across N levels, implying that GY improvement in pearl millet could be possible by selecting genotypes that attain a higher NUtE and NHI across N levels. Similarly, a significant correlation between GY and NUtE was reported in maize, wheat, and oilseed rape (He et al., 2017; Belete et al., 2018). Fageria et al. (2010) reported that strong associations between GY and NUE related traits could be a better option for GY improvements under limited N condition. Furthermore, GY was negatively correlated with NPG and NPS across N levels. Previous studies also reported a similar significant negative correlation between grain N concentration and GY (Sinebo et al., 2004). In the present study, a strong positive correlation of DSY with NUpS and TNUp across N levels was observed indicating that more DSY was with higher NUpS and TNUp, where it was negatively correlated with NPS, NUtE, and NHI. The inverse relationship between NPS and DSY was reported in rice (Subudhi et al., 2020). Identified genetic stocks will be further utilized to carry an in-depth investigation to understand the genes and associated pathways related to NUE. Overall, important correlations identified from this study, will certainly help for the future pearl millet NUE breeding programs.

CONCLUSION

Improving NUE of pearl millet is pivotal for sustainable crop growth and yield especially under low nitrogen soils. Likewise, improving crop productivity using N fertilization is important for achieving climate resilience. Nevertheless, the genetic improvement of pearl millet NUE depends on the nature and extent of variation among the germplasm. As studies on pearl millet NUE are still nascent, this study was aimed to derive morphologic and agronomical traits associated with NUE in a

set of 380 diverse lines under different N levels. The first step in this research revealed extensive genetic variations across the diversity panel and mapping population parents under N_0 and N_{100} conditions. Also, large environmental variations (including N inputs) for GY and its related NUE traits were observed. Nitrogen limitation resulted in the reduction of GY and DSY in N_0 , as compared with N_{50} and N_{100} . Under low N, the best grain yielding genotypes were identified and considered as nitrogen-insensitive (NIS). Genotypes IP10820, IP17720, ICMB01222-P1, IP10379, ICMB89111-P2, IP8069, ICMB90111-P2, ICMV-IS89305, and ICMV221 (ICMV88904) proved to be the most efficient genotypes in terms of GY at low and high N levels, and indeed shows their inherent genotypic plasticity toward N application. Furthermore, the use of NIS genotypes will help in breeding N-efficient genotypes for arid and marginal agro-ecologies. Overall results suggest that genotypes with more yield and high to moderate NUE can be used as parents for the breeding of N efficient genotypes. The lines identified from the present study, coupled with multi-omics technologies can help to identify candidate genes for NUE in pearl millet. The available genetic stocks will be useful to carry an in-depth dissection of the genes and pathways related to NUE. These contrasting genetic stocks may help to map the genome, transcriptome, proteome, and metabolome signatures of the traits leading to a better understanding of NUE in pearl millet. These current findings may also help the farmers for optimizing the use of fertilizer inputs for economic and environmentally sustainable food production.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

RKS and RG planned and designed this research. VP, MP, SB, RR, and VT performed the field experiments, phenotyping, and

analyzed the data. AR, RD, and GA carried out statistical analysis. RKS and RG interpreted the data. VP, MP, RKS, and RG wrote the manuscript. All the authors have made their contribution in editing the manuscript for publication.

FUNDING

Funding support from the Department of Biotechnology (Grant# BT-IN-UK-VNC-42-RG-2015-16), Government of India, the Biotechnology and Biological Sciences Research Council (GCRF/BBSRC Grant # BB/P027970/1), United Kingdom, and the CGIAR Research Program on Grain Legumes and Dryland Cereals (CRP-GLDC) is gratefully acknowledged.

ACKNOWLEDGMENTS

We are grateful to Nagaraj Goud N., Ravinder T., and Shyam U. for data collection at the experimental site. We are also thankful to soil science laboratory technicians for Nitrogen (N) analysis in different samples and tissues.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.625915/full#supplementary-material>

Supplementary Table 1 | Details of soil properties before establishment of experimental plot.

Supplementary Table 2 | Phenotypic data of 380 pearl millet genotypes for five traits (SPAD, Leaf number, Leaf area, Grain yield, and Dry stover yield) at three N levels. Mean values of over three seasons (summer 2017, rainy 2017, and summer 2018).

Supplementary Table 3 | Phenotypic data of 380 pearl millet genotypes for six traits (1. N percent in grain 2. N percent in stover 3. N uptake in stover 4. Total N uptake 5. Nitrogen utilization efficiency 6. Nitrogen Harvest Index) at three N levels. Mean values of over two seasons (summer 2017 and summer 2018).

Supplementary Table 4 | Agronomy (days to 50% flowering and plant height) data provided for top five NIS genotypes under N_0 .

REFERENCES

- Alagarswamy, G., and Biding, F. R. (1987). "Genotypic variation in biomass production and nitrogen use efficiency in pearl millet [*Pennisetum americanum* (L.) Leeke]," in *Genetic Aspects of Plant Mineral Nutrition*, eds W. H. Gabelman and B. C. Loughman (Dordrecht: Springer), 281–286. doi: 10.1007/978-94-009-3581-5_25
- Anbessa, Y., Juskiw, P., Good, A., Nyachiro, J., and Helm, J. (2009). Genetic variability in nitrogen use efficiency of spring barley. *Crop Sci.* 49, 1259–1269. doi: 10.2135/cropsci2008.09.0566
- Arduini, I., Masoni, A., Ercoli, L., and Mariotti, M. (2006). Grain yield, and dry matter and nitrogen accumulation and remobilization in durum wheat as affected by variety and seeding rate. *Eur. J. Agron.* 25, 309–318. doi: 10.1016/j.eja.2006.06.009
- Bandyopadhyay, T., Swarbrick, S. M., Jaiswal, V., Gupta, R., Bentley, A. R., Griffiths, H., et al. (2020). Grain number and genotype drive nitrogen-dependent yield response in the C4 model *Setaria italica* (L.) P. Beauv. *bioRxiv* [Preprint]. doi: 10.1101/2020.03.23.003004
- Basava, R. K., Hash, C. T., Mahendrakar, M. D., Kishor, P. B. K., Satyavathi, C. T., Kumar, S., et al. (2019). Discerning combining ability loci for divergent environments using chromosome segment substitution lines (CSSLs) in pearl millet. *PLoS One* 14:e0218916. doi: 10.1371/journal.pone.0218916
- Belete, F., Dechassa, N., Molla, A., and Tana, T. (2018). Effect of nitrogen fertilizer rates on grain yield and nitrogen uptake and use efficiency of bread wheat (*Triticum aestivum* L.) varieties on the Vertisols of central highlands of Ethiopia. *Agric. Food Secur.* 7, 1–12. doi: 10.1186/s40066-018-0231-z
- Biding, F. R., Blümmel, M., Hash, C. T., and Choudhary, S. (2010). Genetic enhancement for superior food-feed traits in a pearl millet (*Pennisetum glaucum* (L.) R. Br.) variety by recurrent selection. *Anim. Nutr. Feed Tech.* 10, 61–68.
- Blümmel, M., Khan, A. A., Vadez, V., Hash, C. T., and Rai, K. N. (2010). Variability in stover quality traits in commercial hybrids of pearl millet (*Pennisetum glaucum* (L.) R. Br.) and grain-stover trait relationships. *Anim. Nutr. Feed Tech.* 10, 29–38.
- Bouchet, A. S., Nesi, N., Bissuel, C., Bregeon, M., Laripe, A., Navier, H., et al. (2014). Genetic control of yield and yield components in winter oilseed rape

- (*Brassica napus* L.) grown under nitrogen limitation. *Euphytica* 199, 183–205. doi: 10.1007/s10681-014-1130-4
- Chan-Navarrete, R., Kawai, A., Dolstra, O., Van Bueren, E. T. L., and Van der Linden, C. G. (2014). Genetic diversity for nitrogen use efficiency in spinach (*Spinacia oleracea* L.) cultivars using the Ingstad model on hydroponics. *Euphytica* 199, 155–166. doi: 10.1007/s10681-014-1186-1
- Chen, B., Xu, K., Li, J., Li, F., Qiao, J., Li, H., et al. (2014). Evaluation of yield and agronomic traits and their genetic variation in 488 global collections of *Brassica napus* L. *Genet. Resour. Crop Evol.* 61, 979–999. doi: 10.1007/s10722-014-0091-8
- D'Andrea, A. C., and Casey, J. (2002). Pearl millet and Kintampo subsistence. *Afr. Archaeol. Rev.* 19, 147–173. doi: 10.1023/A:1016518919072
- Ertiro, B. T., Beyene, Y., Das, B., Mugo, S., Olsen, M., Oikeh, S., et al. (2017). Combining ability and testcross performance of drought-tolerant maize inbred lines under stress and non-stress environments in Kenya. *Plant Breed.* 136, 197–205. doi: 10.1111/pbr.12464
- Fageria, N. K. (2014). Nitrogen harvest index and its association with crop yields. *J. Plant Nutr.* 37, 795–810. doi: 10.1080/01904167.2014.881855
- Fageria, N. K., and Baligar, V. C. (2003). Methodology for evaluation of lowland rice genotypes for nitrogen use efficiency. *J. Plant Nutr.* 26, 1315–1333. doi: 10.1081/PLN-120020373
- Fageria, N. K., Baligar, V. C., and Li, Y. C. (2008). The role of nutrient efficient plants in improving crop yields in the twenty first century. *J. Plant Nutr.* 31, 1121–1157. doi: 10.1080/01904160802116068
- Fageria, N. K., De Moraes, O. P., and Dos Santos, A. B. (2010). Nitrogen use efficiency in upland rice genotypes. *J. Plant Nutr.* 33, 1696–1711. doi: 10.1080/01904167.2010.496892
- FAO (2019). FAO World Fertilizer Trends and Outlook to 2020. Available online at: <http://www.fao.org/3/a-i6895e.pdf> (accessed June 11, 2019).
- Gianquinto, G., Goffart, J. P., Olivier, M., Guarda, G., Colauzzi, M., Dalla Costa, L., et al. (2004). The use of hand-held chlorophyll meters as a tool to assess the nitrogen status and to guide nitrogen fertilization of potato crop. *Potato Res.* 47, 35–80. doi: 10.1007/bf02731970
- Good, A. G., Shrawat, A. K., and Muench, D. G. (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? *Trends Plant Sci.* 9, 597–605. doi: 10.1016/j.tplants.2004.10.008
- Govintharaj, P., Gupta, S. K., Maheswaran, M., Sumathi, P., and Atkari, D. G. (2018). Correlation and path coefficient analysis of biomass yield and quality traits in forage type hybrid parents of pearl millet. *Int. J. Pure Appl. Biosci.* 6, 1056–1061. doi: 10.18782/2320-7051.5992
- Gupta, S. K., Nepolean, T., Sankar, S. M., Rathore, A., Das, R. R., Rai, K. N., et al. (2015). Patterns of molecular diversity in current and previously developed hybrid parents of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Am. J. Plant Sci.* 6, 1697–1712. doi: 10.4236/ajps.2015.611169
- Hakeem, K. R., Chandna, R., Ahmad, A., Qureshi, M. I., and Iqbal, M. (2012). Proteomic analysis for low and high nitrogen-responsive proteins in the leaves of rice genotypes grown at three nitrogen levels. *Appl. Biochem.* 168, 834–850. doi: 10.1007/s12010-012-9823-4
- Hawkesford, M. J. (2017). Genetic variation in traits for nitrogen use efficiency in wheat. *J. Exp. Bot.* 68, 2627–2632. doi: 10.1093/jxb/erx079
- Hawkesford, M. J., and Griffiths, S. (2019). Exploiting genetic variation in nitrogen use efficiency for cereal crop improvement. *Curr. Opin.* 49, 35–42. doi: 10.1016/j.pbi.2019.05.003
- He, H., Yang, R., Li, Y., Ma, A., Cao, L., Wu, X., et al. (2017). Genotypic variation in nitrogen utilization efficiency of oilseed rape (*Brassica napus*) under contrasting N supply in pot and field experiments. *Front. Plant Sci.* 8:1825. doi: 10.3389/fpls.2017.01825
- Hickman, J. E., Palm, C. A., Mutuo, P., Melillo, J. M., and Tang, J. (2014). Nitrous oxide (N₂O) emissions in response to increasing fertilizer addition in maize (*Zea mays* L.) agriculture in western Kenya. *Nutr. Cycling Agroecosyst.* 100, 177–187. doi: 10.1007/s10705-014-9636-7
- Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *J. Exp. Bot.* 58, 2369–2387. doi: 10.1093/jxb/erm097
- Inthapanya, P., Sihavong, P., Sihathap, V., Chanphengsay, M., Fukai, S., and Basnayake, J. (2000). Genotype differences in nutrient uptake and utilisation for grain yield production of rainfed lowland rice under fertilized and non-fertilized conditions. *Field Crops Res.* 65, 57–68. doi: 10.1016/S0378-4290(99)00070-2
- Kiran, T. V., Vijayalakshmi, P., Rao, Y. V., Swamy, K. N., Kondamudi, R., Srikanth, B., et al. (2016). Effects of nitrogen limitation on antioxidant enzymes, chlorophyll content and grain yield of rice genotypes. *Asian Res. J. Agri.* 1, 1–10. doi: 10.9734/arja/2016/28503
- Kitajima, K., and Hogan, K. P. (2003). Increases of chlorophyll a/b ratios during acclimation of tropical woody seedlings to nitrogen limitation and high light. *Plant Cell Environ.* 26, 857–865. doi: 10.1046/j.1365-3040.2003.01017.x
- Ladha, J. K., Kirk, G. J. D., Bennett, J., Peng, S., Reddy, C. K., Reddy, P. M., et al. (1998). Opportunities for increased nitrogen-use efficiency from improved lowland rice germplasm. *Field Crops Res.* 56, 41–71. doi: 10.1016/S0378-4290(97)00123-8
- Le Gouis, J., Béghin, D., Heumez, E., and Pluchard, P. (2000). Genetic differences for nitrogen uptake and nitrogen utilization efficiencies in winter wheat. *Eur. J. Agron.* 12, 163–173. doi: 10.1016/S1161-0301(00)00045-9
- Li, F., Ai, T., Zhou, S., Nie, J., and Liu, F. (2003). Influence of slow-release nitrogen fertilizers on lowland rice yield and nitrogen use efficiency. *Chin. J. Soil Sci.* 35, 311–315.
- Li, P., Chen, F., Cai, H., Liu, J., Pan, Q., Liu, Z., et al. (2015). A genetic relationship between nitrogen use efficiency and seedling root traits in maize as revealed by QTL analysis. *J. Exp. Bot.* 66, 3175–3188. doi: 10.1093/jxb/erv127
- Long, J. R., Wan, Y. Z., Song, C. F., Jian, S. U. N., and Qin, R. J. (2013). Effects of nitrogen fertilizer level on chlorophyll fluorescence characteristics in flag leaf of super hybrid rice at late growth stage. *Rice Sci.* 20, 220–228. doi: 10.1016/s1672-6308(13)60138-9
- Lopez-Bellido, R. J., Shepherd, C. E., and Barraclough, P. B. (2004). Predicting post-anthesis N requirements of bread wheat with a Minolta SPAD meter. *Eur. J. Agron.* 20, 313–320. doi: 10.1016/S1161-0301(03)00025-X
- Ma, B. L., Biswas, D. K., Herath, A. W., Whalen, J. K., Ruan, S. Q., Caldwell, C., et al. (2015). Growth, yield, and yield components of canola as affected by nitrogen, sulfur, and boron application. *J. Plant Nutr. Soil Sci.* 178, 658–670. doi: 10.1002/jpln.201400280
- Maman, N., Mason, S. C., and Lyon, D. J. (2006). Nitrogen rate influence on pearl millet yield, nitrogen uptake, and nitrogen use efficiency in Nebraska. *Commun. Soil Sci. Plant Anal.* 37, 127–141. doi: 10.1080/00103620500406112
- Manning, K., Pelling, R., Higham, T., Schwenniger, J. L., and Fuller, D. Q. (2011). 4500-Year old domesticated pearl millet (*Pennisetum glaucum*) from the Tilemsi Valley, Mali: new insights into an alternative cereal domestication pathway. *J. Archaeol. Sci.* 38, 312–322. doi: 10.1038/nrg2612
- Mhaskar, N. V., Dongale, J. H., Dademal, A. A., and Khanvilkar, S. A. (2005). Performance of scented rice varieties under different nitrogen levels in lateritic soil of Konkan. *Oryza* 42:323.
- Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. *J. Agron.* 74, 562–564. doi: 10.2134/agronj1982.00021962007400030037x
- Monostori, I., Árendás, T., Hoffman, B., Galiba, G., Gierczik, K., Szira, F., et al. (2016). Relationship between SPAD value and grain yield can be affected by cultivar, environment and soil nitrogen content in wheat. *Euphytica* 211, 103–112. doi: 10.1007/s10681-016-1741-z
- Monostori, I., Szira, F., Tondelli, A., Árendás, T., Gierczik, K., Cattivelli, L., et al. (2017). Genome-wide association study and genetic diversity analysis on nitrogen use efficiency in a Central European winter wheat (*Triticum aestivum* L.) collection. *PLoS One* 12:e0189265. doi: 10.1371/journal.pone.0189265
- Muchow, R. C. (1998). Nitrogen utilization efficiency in maize and grain sorghum. *Field Crops Res.* 56, 209–216. doi: 10.1016/S0378-4290(97)00132-9
- Namai, S., Toriyama, K., and Fukuta, Y. (2009). Genetic variations in dry matter production and physiological nitrogen use efficiency in rice (*Oryza sativa* L.) varieties. *Breed. Sci.* 59, 269–276. doi: 10.1270/jsbbs.59.269
- Pramanik, K., and Bera, A. K. (2013). Effect of seedling age and nitrogen fertilizer on growth, chlorophyll content, yield and economics of hybrid rice (*Oryza sativa* L.). *Int. J. Agron. Plant Prod.* 4, 3489–3499.
- Presterl, T., Seitz, G., Landbeck, M., Thient, E. M., Schmidt, W., and Geiger, H. H. (2003). Improving nitrogen-use efficiency in European maize. *Crop Sci.* 43, 1259–1265. doi: 10.2135/cropsci2003.1259
- Rai, K. N., Blummel, M., Singh, A. K., and Rao, A. S. (2012). Variability and relationships among forage yield and quality traits in pearl millet. *Eur. J. Plant Sci. Biotechnol.* 6, 118–124.

- Rao, I. S., Neeraja, C. N., Srikanth, B., Subrahmanyam, D., Swamy, K. N., Rajesh, K., et al. (2018). Identification of rice landraces with promising yield and the associated genomic regions under low nitrogen. *Sci. Rep.* 8:9200. doi: 10.1038/s41598-018-27484-0
- Raun, W. R., Johnson, G. V., and Westerman, R. L. (1999). Fertilizer nitrogen recovery in long-term continuous winter wheat. *Soil Sci. Soc. Am. J.* 63, 645–650. doi: 10.2136/sssaj1999.03615995006300030030x
- Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K., et al. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. *J. Agron.* 94, 815–820. doi: 10.2134/agronj2002.8150
- Rengel, Z., and Graham, R. D. (1995). Wheat genotypes differ in Zn efficiency when grown in chelate-buffered nutrient solution. *Plant Soil* 176, 307–316. doi: 10.1007/BF00011795
- Russo, T. A., Tully, K., Palm, C., and Neill, C. (2017). Leaching losses from Kenyan maize cropland receiving different rates of nitrogen fertilizer. *Nutr. Cycling Agroecosyst.* 108, 195–209. doi: 10.1007/s10705-017-9852-z
- Sahrawat, K. L., Kumar, G. R., and Murthy, K. V. S. (2002). Sulfuric acid–selenium digestion for multi-element analysis in a single plant digest. *Commun. Soil Sci. Plan* 33, 3757–3765. doi: 10.1081/CSS-120015920
- Šarčević, H., Jukić, K., Ikić, I., and Lovrić, A. (2014). Estimation of quantitative genetic parameters for grain yield and quality in winter wheat under high and low nitrogen fertilization. *Euphytica* 199, 57–67. doi: 10.1007/s10681-014-1154-9
- Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., and Taunk, J. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS One* 10:e0122165. doi: 10.1371/journal.pone.0122165
- Sial, M. A., Arain, M. A., Khanzada, S. H. A. M. A. D. A. D., Naqvi, M. H., Dahot, M. U., and Nizamani, N. A. (2005). Yield and quality parameters of wheat genotypes as affected by sowing dates and high temperature stress. *Pak. J. Bot.* 37, 575–584.
- Sinebo, W., Gretzmacher, R., and Edelbauer, A. (2004). Genotypic variation for nitrogen use efficiency in Ethiopian barley. *Field Crops Res.* 85, 43–60. doi: 10.1016/S0378-4290(03)00135-7
- Singh, T., Shivay, Y. S., and Singh, S. (2004). Effect of date of transplanting and nitrogen on productivity and nitrogen use indices in hybrid and non-hybrid aromatic rice. *Acta Agron. Hung.* 52, 245–252. doi: 10.1556/aagr.52.2004.3.5
- Srikanth, B., Rao, I. S., Surekha, K., Subrahmanyam, D., Voleti, S. R., and Neeraja, C. N. (2016). Enhanced expression of OsSPL14 gene and its association with yield components in rice (*Oryza sativa*) under low nitrogen conditions. *Gene* 576, 441–450. doi: 10.1016/j.gene.2015.10.062
- Srivastava, R., Singh, R. B., Lakshmi, P. V., Srikanth, B., Madhu, P., Satyavathi, C. T., et al. (2019). Genome-wide association studies (GWAS) and genomic selection (GS) in pearl millet: advances and prospects. *Front. Genet.* 10:1389. doi: 10.3389/fgene.2019.01389
- Stahl, A., Friedt, W., Wittkop, B., and Snowden, R. J. (2016). Complementary diversity for nitrogen uptake and utilisation efficiency reveals broad potential for increased sustainability of oilseed rape production. *Plant Soil* 400, 245–262. doi: 10.1007/s11104-015-2726-8
- Subudhi, H. N., Prasad, K. V. S. V., Ramakrishna, C., Rameswar, P. S., Pathak, H., Ravi, D., et al. (2020). Genetic variation for grain yield, straw yield and straw quality traits in 132 diverse rice varieties released for different ecologies such as upland, lowland, irrigated and salinity prone areas in India. *Field Crops Res.* 245:107626. doi: 10.1016/j.fcr.2019.107626
- Tako, E., Reed, S. M., Budiman, J., Hart, J. J., and Glahn, R. P. (2015). Higher iron pearl millet (*Pennisetum glaucum* L.) 967 provides more absorbable iron that is limited by increased polyphenolic content. *Nutr. J.* 14, 1–9. . https://doi: 10.1186/1475-2891-14-11,
- Ulas, A., Behrens, T., Wiesler, F., and Horst, W. J. (2013). Does genotypic variation in nitrogen remobilisation efficiency contribute to nitrogen efficiency of winter oilseed-rape cultivars (*Brassica napus* L.). *Plant Soil* 371, 463–471. doi: 10.1007/s11104-013-1688-y
- Vijayalakshmi, P., Kiran, T. V., Rao, Y. V., Srikanth, B., Rao, I. S., Sailaja, B., et al. (2013). Physiological approaches for increasing nitrogen use efficiency in rice. *Indian J. Plant Physiol.* 18, 208–222. doi: 10.1007/s40502-013-0042-y
- Vijayalakshmi, P., Vishnukiran, T., Kumari, B. R., Srikanth, B., Rao, I. S., Swamy, K. N., et al. (2015). Biochemical and physiological characterization for nitrogen use efficiency in aromatic rice genotypes. *Field Crops Res.* 179, 132–143. doi: 10.1016/j.fcr.2015.04.012
- VSN International (2019). *Genstat for Windows*, 20th Edn. Hemel Hempstead: VSN International.
- Worku, M., Bänziger, M., Erley, G. S. A. M., Friesen, D., Diallo, A. O., and Horst, W. J. (2007). Nitrogen uptake and utilization in contrasting nitrogen efficient tropical maize hybrids. *Crop Sci.* 47, 519–528. doi: 10.2135/cropsci2005.05.0070
- Würschum, T. (2012). Mapping QTL for agronomic traits in breeding populations. *Theor. Appl. Genet.* 125, 201–210. doi: 10.1007/s00122-012-1887-6
- Yang, H., Yang, J., Lv, Y., and He, J. (2014). SPAD values and nitrogen nutrition index for the evaluation of rice nitrogen status. *Plant Pro. Sci.* 17, 81–92. doi: 10.1626/ppp.17.81

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Pujarula, Pusuluri, Bollam, Das, Ratnala, Adapala, Thuraga, Rathore, Srivastava and Gupta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Genetic Gains in Pearl Millet in India: Insights Into Historic Breeding Strategies and Future Perspective

Om Parkash Yadav^{1*}, S. K. Gupta², Mahalingam Govindaraj², Rajan Sharma²,
Rajeev K. Varshney^{2,3}, Rakesh K. Srivastava², A. Rathore² and Rajendra Singh Mahala⁴

¹ ICAR-Central Arid Zone Research Institute, Jodhpur, India, ² International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, ³ State Agricultural Biotechnology Centre, Centre for Crop and Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA, Australia, ⁴ SeedWorks International Pvt Ltd, Hyderabad, India

OPEN ACCESS

Edited by:

Petr Smýkal,
Palacký University, Olomouc, Czechia

Reviewed by:

Flavio Breseghello,
Brazilian Agricultural Research
Corporation (EMBRAPA), Brazil
Michael Benjamin Kantar,
University of Hawaii, United States

*Correspondence:

Om Parkash Yadav
opyadav21@yahoo.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 22 December 2020

Accepted: 19 February 2021

Published: 30 March 2021

Citation:

Yadav OP, Gupta SK, Govindaraj M,
Sharma R, Varshney RK,
Srivastava RK, Rathore A and
Mahala RS (2021) Genetic Gains in
Pearl Millet in India: Insights Into
Historic Breeding Strategies and
Future Perspective.
Front. Plant Sci. 12:645038.
doi: 10.3389/fpls.2021.645038

Pearl millet (*Pennisetum glaucum* R. Br.) is an important staple and nutritious food crop in the semiarid and arid ecologies of South Asia (SA) and Sub-Saharan Africa (SSA). In view of climate change, depleting water resources, and widespread malnutrition, there is a need to accelerate the rate of genetic gains in pearl millet productivity. This review discusses past strategies and future approaches to accelerate genetic gains to meet future demand. Pearl millet breeding in India has historically evolved very comprehensively from open-pollinated varieties development to hybrid breeding. Availability of stable cytoplasmic male sterility system with adequate restorers and strategic use of genetic resources from India and SSA laid the strong foundation of hybrid breeding. Genetic and cytoplasmic diversification of hybrid parental lines, periodic replacement of hybrids, and breeding disease-resistant and stress-tolerant cultivars have been areas of very high priority. As a result, an annual yield increase of 4% has been realized in the last three decades. There is considerable scope to further accelerate the efforts on hybrid breeding for drought-prone areas in SA and SSA. Heterotic grouping of hybrid parental lines is essential to sustain long-term genetic gains. Time is now ripe for mainstreaming of the nutritional traits improvement in pearl millet breeding programs. New opportunities are emerging to improve the efficiency and precision of breeding. Development and application of high-throughput genomic tools, speed breeding, and precision phenotyping protocols need to be intensified to exploit a huge wealth of native genetic variation available in pearl millet to accelerate the genetic gains.

Keywords: pearl millet, hybrid breeding, genetic gain, disease resistance, drought tolerance, biofortification, heat tolerance

INTRODUCTION

Pearl millet (*Pennisetum glaucum* R. Br.) is an important crop in the semiarid and arid ecologies of South Asia (SA) and sub-Saharan Africa (SSA) that are characteristically challenged by low and erratic rainfall and high mean temperature and simultaneously have soils with low organic carbon and poor water-holding capacity (Serba et al., 2020). Besides its unmatched capacity to tolerate drought, pearl millet has also built-in adaptation to soils with low fertility. Because of its remarkable ability to respond to favorable environments because of its short developmental stages, high photosynthetic efficiency, and abundant capacity for high growth rate, pearl millet is

an excellent crop for the short growing season and under improved crop management (Yadav and Rai, 2013) and is emerging as an important alternative crop for feed, food, fodder, and relay crop in Brazil, Canada, Mexico, United States, West Asia and North Africa region, and Central Asia.

Pearl millet is valued for its nutrient-rich grain for human consumption and its green fodder and dry stover for livestock (Andrews and Kumar, 1992) and forms the basis of livelihood and nutritional security for more than 90 million people in SSA and SA (Serba et al., 2020). Pearl millet demand is anticipated to increase in the future because of increasing human and livestock populations in SSA and SA and as a healthy food and other industrial uses (Rai et al., 2008). Its cultivation may further extend in the areas where maize and sorghum are cultivated because of depleting water resources. Pearl millet production is likely to become more challenging because of predicted intense drought stress, rise in temperature, and greater disease incidences in SSA (Sultan et al., 2013) and SA (Rama Rao et al., 2019). Therefore, its production must be increased at a much faster rate and more so in challenging agroecologies. Increasing production must come through enhancement in productivity as there is little scope to enhance production by expanding its cultivation, especially in SA.

Enormous progress has been made in India to improve productivity by developing high-yielding cultivars and their improved agronomic management during the last six decades (Jukanti et al., 2016). The accomplishments of pearl millet breeding are often referred to as one of the greatest success stories in Indian agriculture (Yadav et al., 2019). However, the biological potential of pearl millet has not been fully realized as indicated by the current 1.2 ton/ha national productivity in comparison to the productivity level of 4–5 tons/ha in the summer season in northwestern India. An attempt is made here to critically analyze past breeding strategies followed in pearl millet. We also examine the prospects of further accelerating genetic gains to meet the greater demand for pearl millet and to make its cultivation more profitable.

HISTORIC BREEDING STRATEGIES

Breeding strategies in pearl millet have evolved very comprehensively over several decades, taking into account understanding of its pollination behavior, challenges in its production, access to germplasm, and accumulated knowledge in the fields of its genetics, physiology, pathology, and so on.

Pearl millet is a highly cross-pollinated species, with outcrossing rates of more than 85%, because of its protogynous nature of flowering. Therefore, individual plants of natural populations mate randomly and are highly heterozygous and heterogeneous. Early breeding efforts in genetic improvement of pearl millet, which started as early as the 1930s, attempted to capitalize on such existing genetic variation within traditional landraces by subjecting them to simple mass selection (Athwal, 1961). The greater urgency for population improvement programs started with the acquisition of a diverse range of germplasm from across the world in the 1970s with the

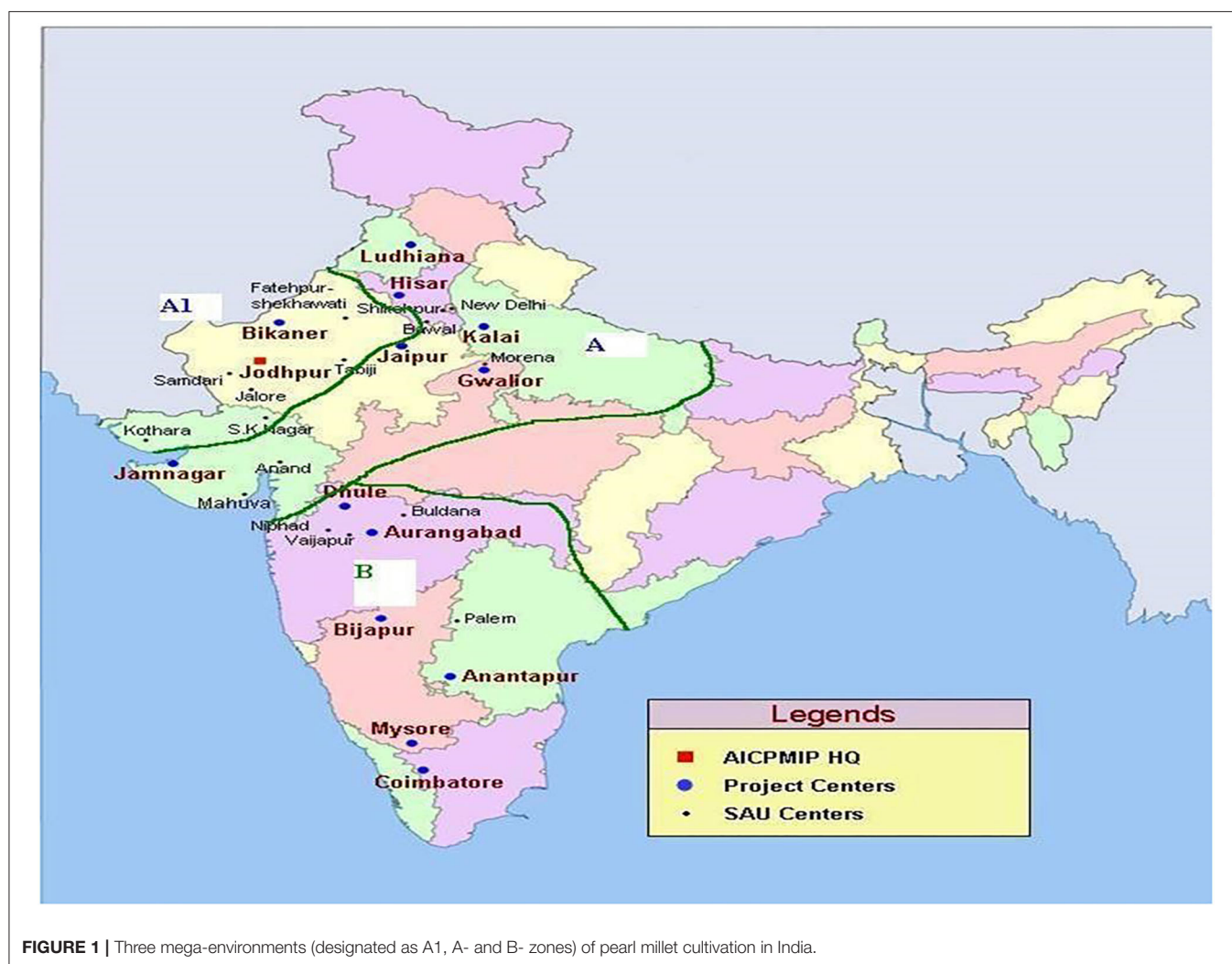
establishment of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) (Gill, 1991; Witcombe, 1999). Eventually, a large number of populations and trait-based composites of the broad genetic base were established, and a diverse range of elite breeding materials was developed (Rai and Anand Kumar, 1994; Rai et al., 2006).

Like maize, the heterosis was also observed in pearl millet for grain yield (Athwal, 1966), but the hermaphrodite nature of flowers of small size limited the ability to exploit it at the commercial level. There were some innovative attempts in the 1950s to exploit heterosis through developing “chance hybrids” that included growing two parental populations of similar maturity in the mixture and allowing them to cross-pollinate to produce seed that contained ~40% hybrid seed (Gill, 1991). The chance hybrids outyielded local varieties by 10–15% but could not become popular because of lack of efficient seed production programs and their limited genetic superiority.

The discovery of cytoplasmic male sterility (CMS) in 1958 (Burton, 1965) at Tifton, Georgia, and the availability of good fertility restorers in Indian germplasm led to the development of commercial hybrids, making a quick and significant impact on pearl millet production (Gill, 1991). However, the cultivation of a limited number of hybrids over several years led to a downy mildew (DM) epidemic in the early 1970s (Safeulla, 1976). Such epidemics reappeared whenever a few hybrids occupied a large area year after year (Singh, 1995). No association of A1 cytoplasm was established with the DM epidemic (Yadav et al., 1993; Yadav, 1996).

Recurring DM epidemics in pearl millet hybrids in India prompted to intensify efforts on genetic diversification of hybrid parental lines, especially after the 1980s. This involved both cytoplasmic and nuclear diversification of parental lines. In addition to A1 CMS source (Burton, 1965), A2, A3 (Athwal, 1961, 1966), A4 (Hanna, 1989), and A5 sources (Rai, 1995) were reported. Extensive characterization of these sources established instability of A2 and A3 sources, whereas A4 and A5 were found as more promising (Rai et al., 1996). This was followed by the development and dissemination of 89 A lines based on A4 and 27 based on A5 source by ICRISAT, but utilization of these two sources remained restricted because of lack of suitable restorers. ICRISAT initiated breeding efforts for developing restorers, especially for the A5 CMS system (Rai et al., 1996, 2009b). Research programs in India have now started breeding both A and R lines and developing hybrids based on these CMS systems. The understanding of the genetics of A4 (Gupta et al., 2012a) and A5 CMS (Gupta et al., 2018) helped in the well-organized and efficient utilization of these CMS sources.

A range of germplasm material from India and Africa with diverse phenotypic characteristics, such as tillering, panicle size, earliness, grain size, grain color, and so on, was strategically exploited to diversify the genetic base of both seed and restorer parents (Andrews and Anand Kumar, 1996; Rai et al., 2009a; Yadav et al., 2012c; Patil et al., 2020). In the last four decades, hybrid breeding has received a very high priority in India using genetically diverse parental lines targeting various production ecologies that have helped to intensify the genetic gains (Rai et al., 2009a; Yadav et al., 2012a).



TRAIT PRIORITIZATION

Pearl millet is cultivated under diverse agroecologies ranging from near-optimum environments (with high use of irrigation and chemical fertilizers) to extremely challenging drought-prone environments (with little external inputs). This led to the prioritization of research to improve necessary phenotypic traits, climate adaption, disease resistance, and nutritional traits taking full cognizance of production constraints and differential requirements of various regions.

Phenotypic and Productivity-Related Traits and Their Linkage With Megaenvironments

A single pearl millet cultivar cannot be expected to perform well under all the environmental conditions, and a cultivar planted outside its adaptation zone would suffer yield reduction due to significant genotype \times environment interactions. Therefore, breeding and evaluation require a subdivision of the testing environments into relatively more homogeneous groups of

locations, called megaenvironments, where specific genotypes can be targeted for individual megaenvironment.

Indian pearl millet cultivation area has been divided into three megaenvironments (designated as A₁, A, and B zones) considering the geographical location, rainfall pattern, local adaptation, and other environmental conditions (Gupta et al., 2013). The A zone consists of parts of northern India receiving >400 mm of annual rainfall (**Figure 1**). The A₁ zone consists of parts of northwestern India receiving <400 mm of annual rainfall, whereas the B zone accounts for the area in peninsular India receiving more than >400-mm annual rainfall. At present, ~75% of the pearl millet is grown in A and A₁ zones and 25% in B zone. Different pearl millet breeding programs in India have developed their product profiles, depending on the need of their target megaenvironment. ICRISAT, as a CGIAR center, which has a global mandate, is targeting all the megaenvironments to support national breeding programs.

High grain yield, disease resistance, and maturity duration of 75–85 days, as per the agroecological requirements, have been accorded the highest priority (Yadav and Rai, 2013). Because of

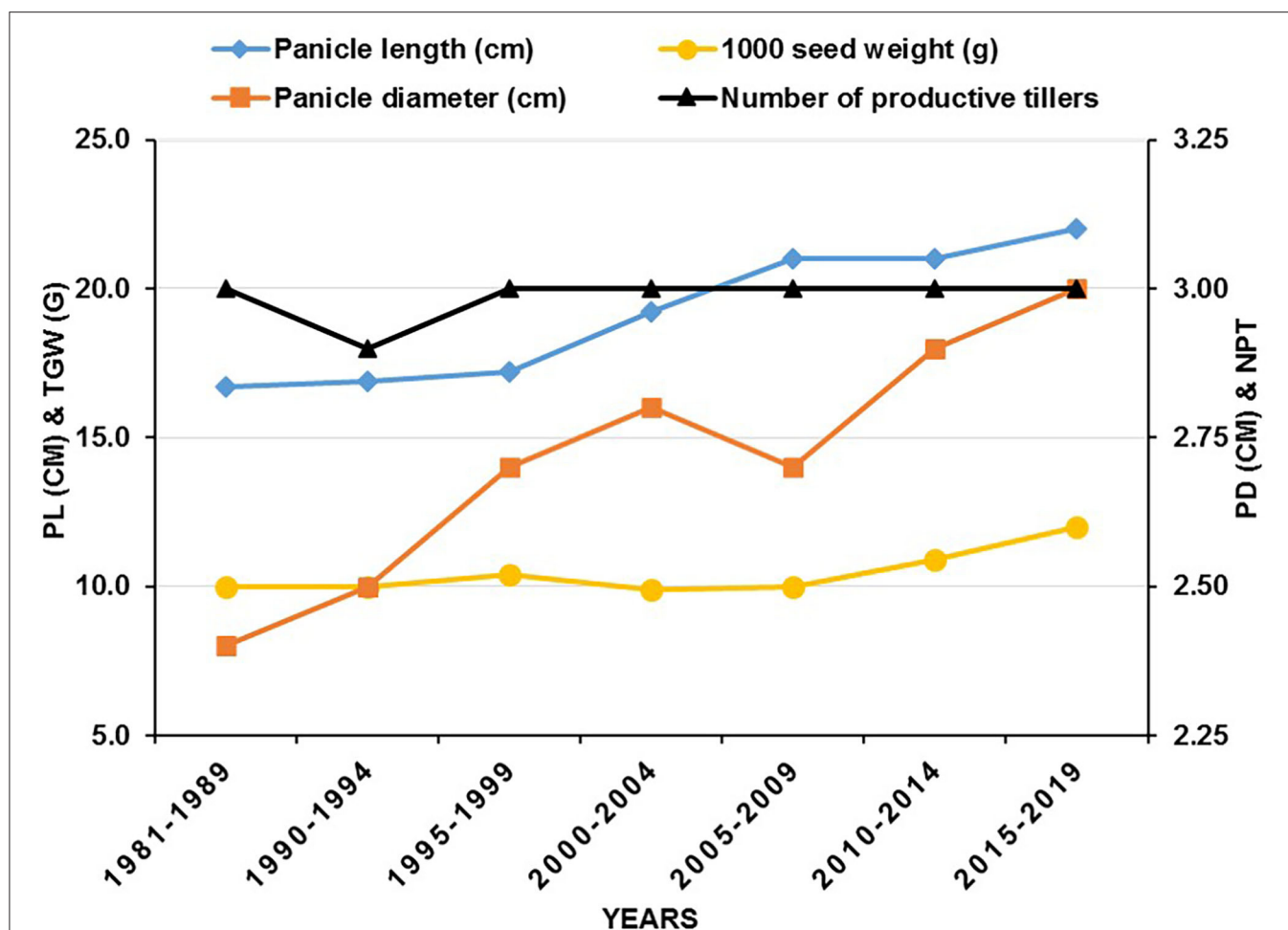


FIGURE 2 | Comparison of different traits in seed parents (A/B-lines) developed between 1981 and 2019 at ICRISAT, Patancheru (PL, panicle length; TGW, thousand grains weight; PD, panicle diameter; NPT, number of productive tillers per plant).

the growing importance of dry stover for fodder purposes, there has been a considerable emphasis, in recent times, on breeding for dual-purpose cultivars producing both high stover and grain yields (Yadav et al., 2012b). Some of the traits in hybrids common to all the environments include lodging resistance, compact panicles, good exertion, and good seed set.

Traits that have regional preferences include various maturity types, tillering, panicle size (a combination of panicle length and girth), and seed size and have been strategically manipulated. Genetic variation for these traits is widely available with high heritability (Yadav et al., 2004, 2017), and simple selection has been, therefore, very successful (Rattunde et al., 1989). Panicle length has increased from 16.7 to 22.0 cm, panicle diameter from 2.4 to 3.0 cm, and 1,000-grain weight from 10 to 12 g in seed parents developed during the last four decades (Figure 2). However, the mean productive tillers per plant were found unchanged.

The d_2 dwarfing gene-based shorter height is the most dominant plant type developed in seed parents breeding (Rai and Hanna, 1990; Rai and Rao, 1991) as it reduces the risk of lodging

in high-management conditions and helps in easy detection of off-type and pollen shedders in the seed production plots.

The A lines have been bred for complete and stable male sterility and B lines for profuse pollen production ability across the seasons and sites. In the breeding of A lines, high grain yield potential, both as lines *per se* and in hybrids (i.e., combining ability), is the most important consideration. High yield, however, is achieved in combination with other agronomic and farmers' preferred traits. The foremost requirement in the restorer lines is to produce highly fertile hybrids and to produce profuse pollen that remains viable at air temperatures as high as 42–44°C. It is desirable to breed pollinators that are taller than A lines, usually in the range of 150–180-cm height.

There has been a clear distinction between the public and private sector hybrid breeding programs regarding trait-based breeding. For instance, private sector breeding programs have largely focused on relatively better-endowed environments (A and B zones), giving greater emphasis to breeding dual-purpose hybrids. As a result, private sector hybrids are generally taller, later in maturity, with longer panicles and less number of effective

tillers/plant [All India Coordinated Research Project on Pearl Millet (AICPMIP), unpublished data] as these traits are reported to contribute toward higher productivity in better-endowed environments. In contrast, most of the public sector hybrids are generally shorter in height and early in maturity, with smaller panicles and a higher number of effective tillers/plant. The private sector has also placed greater emphasis on breeding high grain yield and large-seeded hybrids.

The strategy of trait diversification led to the development of a diverse array of high-yielding hybrids with adequate adaptation to different agroecologies. This was evidenced from a recent investigation where 122 commercial hybrids showed the existence of significant variation for flowering time (42–58 days), tillering (1.1–4.4 panicles/plant), individual grain size (7.6–17.3 mg), plant height (185–268 cm), and panicle length (20–33 cm), which highlighted the successful efforts of the national program of pearl millet improvement toward genetic diversification of hybrids (Yadav et al., 2017).

Climate Adaptation

Pearl millet is challenged by drought and heat stress in various production environments of SA and SSA. While drought is of common occurrence, heat stress assumes importance in specific regions.

Drought Tolerance

Drought, caused by the low rainfall and its erratic distribution, is the primary abiotic production constraint in SA and SSA. Efforts have therefore been made for mapping and delineation of drought-prone regions to define the target population of environments or megaenvironments (Gupta et al., 2013), which are highly variable in terms of their timing, intensity, and duration of drought (van Oosterom et al., 1995).

The subject of drought tolerance in pearl millet has remained, so far, a strategic research issue, and therefore, the response to drought has been studied comprehensively. Drought during the seedling stage results in poor plant stands reducing yield severely (Soman et al., 1987). Drought during the vegetative stage of growth has a little adverse effect on productivity (Bidinger et al., 1987) as there is a significant increase in panicle number, which is a reflection of a compensation mechanism for a damaged main shoot during drought (van Oosterom et al., 2003). However, the grain-filling stage is the most sensitive growth stage to drought as both grain number and grain size are significantly reduced when the crop is exposed to drought stress at this stage (Fussell et al., 1991).

Dissection of drought tolerance in terms of physiology, phenology, and morphology of the crop has led to the understanding of the yield formation process under drought (van Oosterom et al., 2003; Yadav, 2011), helping breeders to identify and target specific traits in different drought environments. Use of physiological traits as selection criteria for drought tolerance has been very challenging especially when dealing with a large number of genotypes in breeding nurseries. Drought escape mechanism has been successfully exploited by targeting early maturity for getting greater genetic gains in the drought-prone regions of northwestern India (Yadav et al., 2011). Morphological

traits such as high tillering, small grain size, and shorter grain filling periods that can be measured easily have been manipulated successfully in breeding programs as there is an abundant variation available for these traits (Yadav et al., 2017).

The role of adapted germplasm has also been emphasized for drought breeding as the measured performance under drought stress is largely a result of adaptation to stress conditions (Yadav et al., 2009, 2012c). Hybridization of adapted landraces with elite genetic material creates new gene combinations that lead to amalgamating of adaptation to stress environments and high productivity (Presterl and Weltzien, 2003; Yadav and Rai, 2011, Patil et al., 2020). Genome regions underlying drought tolerance-related traits have been identified and mapped (Yadav et al., 2002, 2004; Serba and Yadav, 2016). Several such genomic regions are being manipulated to enhance drought tolerance (Bidinger et al., 2007; Sharma et al., 2014).

Heat Tolerance

Optimum temperature for normal growth of pearl millet is 33/34°C. Higher temperatures can affect the pearl millet both at the seedling and reproductive stages. Climate change models project that the pearl millet yield in SSA and SA will decrease by 6–17% by 2050 (Knox et al., 2011).

In India and western and southern Africa, soil surface temperatures often exceed 45°C and may sometimes reach 60°C and are one of the most important factors causing poor plant stands as pearl millet seedlings are most vulnerable to high temperatures during their first 10 days (Soman et al., 1981; Peacock et al., 1993). Therefore, improvement in heat tolerance at the seedling stage assumes importance. Genetic differences in seedling survival under high soil surface temperatures have been reported (Peacock et al., 1993), and selection for greater seedling emergence under artificial screening technique (Soman and Peacock, 1985) has been found effective (Lynch, 1994). A laboratory method based on measuring membrane thermostability has been developed (Howarth et al., 1997) to assess the differences in seedling heat tolerance. There has been no report of manipulating this trait in breeding programs either in SA or SSA in the last three decades.

During the last decade, pearl millet has emerged as a highly productive and remunerative crop in the hot and dry summer season in the northern and western parts of India (Yadav and Rai, 2013). With higher air temperatures (often >42°C) coinciding with flowering in this season, the crop suffers from reproductive sterility, leading to drastic reductions in seed set and finally into lesser grain yield (Gupta et al., 2015b; Djanaguiraman et al., 2018). Heat tolerance as the reproductive stage has emerged as an important target trait to enhance genetic gains.

Flowering-period heat stress screening protocols have been standardized for screening under both controlled environment facilities (growth chambers) and field conditions in heat stress-prone target ecology (Gupta et al., 2015b). Multilocal and multiyear field screening in summer season involving a large number of hybrid parental lines, germplasm accessions, and improved populations established that stigma is more heat-sensitive than pollen; large genetic variation exists between breeding lines and within open-pollinating populations; the

boot-leaf stage is more heat-sensitive than panicle-emergence stage, and heat tolerance behaves as a dominant trait. These screenings led to the identification of heat-tolerant breeding materials that have been used to further enhance heat tolerance (Gupta et al., 2016, 2019) and facilitated the pyramiding of heat tolerance in high-yielding hybrids.

Biotic Stress Resistance

Pearl millet is a hardy crop vulnerable to only fewer diseases and insect-pests compared to other major cereals. However, the diseases that attack pearl millet are capable of causing huge damage. Being a crop grown by resource-poor farmers, diseases and pests are best managed through host plant resistance (HPR) as it does not incur an additional cost.

Diseases

Disease management through HPR involves a sound knowledge of biology and epidemiology of diseases, availability of pure culture of the pathogens, effective inoculation techniques, greenhouse and field screening facilities, appropriate disease rating scales, and availability of resistance sources.

Downy Mildew

DM caused by *Sclerospora graminicola* (Sacc.) J. Schröt is the most important disease causing heavy economic losses in India and Africa. Greenhouse and field screening of a large number of germplasm accessions and breeding lines has led to the identification of several resistance sources (Singh S. D. et al., 1987; Singh et al., 1997), which have been extensively used to develop DM resistant hybrids. The diversified genetic base of hybrids has contributed to the control of widespread DM epidemics.

Regular monitoring of changes in the virulence of pathogens and resistance of hybrids has helped in keeping track of the breakdown of resistance in hybrids. Replacing the susceptible hybrid with its disease-resistant version, created by marker-assisted selection, has also been an effective strategy (Hash et al., 2006).

Blast

The blast or leaf spot disease, caused by *Pyricularia grisea* Sacc. [syn. *Magnaporthe grisea*], has emerged as a very destructive disease of pearl millet in the recent past (Rai et al., 2012). Monitoring of virulence of pathogen populations and screening of genetic resources led to the identification of stable resistant lines to develop blast-resistant hybrid parent lines (Sharma et al., 2013; Yella Goud et al., 2016).

Resistance in pearl millet to Indian isolates of *M. grisea* is governed by a single dominant gene (Gupta et al., 2012b; Singh et al., 2018b), which makes the incorporation of resistance much easier. Molecular markers are also being used to identify Quantitative Trait Loci (QTLs) for blast resistance against prevalent pathotypes. Two major blast resistance QTLs, on LG 4 and LG 7 of pearl millet line 863B-P2, were identified that have been used to improve hybrid parental lines.

Rust

Rust (*Puccinia substriata* var. *indica* Ramachar & Cumm) is generally considered as a disease of less importance in the grain

crop; however, it is of great importance in fodder crop where it reduces both the quantity and quality of the produce. The field screening in the late rainy season under high disease pressure and greenhouse screening led to the identification of stable resistance sources (Singh S. D. et al., 1987; Singh et al., 1990; Sharma R. et al., 2020).

Smut and Ergot

Ergot (*Claviceps fusiformis* Lov.) and smut (*Moesziomyces penicillariae* Bref. Vanky) are important floral diseases, and grain yield losses are proportional to their severity. Both pathogens are soil-borne and infect the host at flowering through stigma. Pollination before pathogen infection prevents infection (Thakur and Williams, 1980). These diseases are more severe in wet weather primarily because of pollen wash. The higher susceptibility of hybrids compared to the open-pollinated varieties is attributed to their greater uniformity in flowering time, rather than their cytoplasm (Yadav et al., 1992; Rai and Thakur, 1996).

Understanding the biology and epidemiology of these diseases helped the development of field screening techniques (Thakur and Rai, 2002). A large number of lines have been evaluated for their reaction to these diseases (Thakur and Rai, 2002; Abraham et al., 2019). Very high levels of ergot resistance in the germplasm accessions were not observed; hence, ergot-resistant lines were developed by intermating less susceptible plants and selecting and rescreening resistant progenies for several generations under high disease pressure (Thakur et al., 1993).

Smut resistance is a dominant trait and easily transferable. However, quantitative resistance involving additive and non-additive gene effects has also been reported (Thakur et al., 2011). A large number of lines have been found as resistant to smut and DM (Thakur et al., 1992, 2011). Currently, multilocal testing across seasons to evaluate the severity of smut and ergot on new experimental cultivars is being done to ensure no smut or ergot susceptible cultivar is released for cultivation.

Insect-Pests

More than 100 insect-pests have been reported to be associated with the pearl millet-based cropping system, but only a few of them are potential pests of any significant economic importance. These are shoot fly (*Atherigona approximata*), stem borers (*Chilo partellus* in India and *Coniesta ignefusalis* in western Africa) and white grubs (*Holotrichia consanguinea*) in India, earhead worms (*Helicoverpa armigera*), gray weevil (*Myloccerus* species), and leaf roller (*Marasmia trapezalis*) (Raghvani et al., 2008). Nature of damage and life cycle of these pests have been studied, and control measures developed. The distribution and damage of insect-pests vary in different regions. Long-term monitoring revealed that no single method of control is effective against any insect. It requires an integrated approach including cultural and chemical control (Sharma and Youm, 1999). The insect-pest incidence on commercial cultivars and experimental test genotypes is closely monitored, and no breeding programs are undertaking insect resistance as target trait.

TABLE 1 | Genetic variability for grain iron and zinc content in germplasm, inbreds, commercial cultivars, and mapping populations of pearl millet.

Genetic materials	No of genotypes	Fe (ppm)	Zn (ppm)	References
Germplasm	191	51–121	46–87	Rai et al., 2014
Inbreds	45	34–102	34–84	Govindaraj et al., 2013
	28	30–82	27–56	Kanatti et al., 2014
	281	35–116	21–80	Pujar et al., 2020
Commercial hybrids	52	47–85	36–70	Velu et al., 2008
	120	46–56	37–44	
Populations (OPVs)	68	42–80	27–50	Velu et al., 2008
	18	42–67	37–52	
Population progenies	240	29–89	32–71	Govindaraj et al., 2016
	299	31–143	35–82	Govindaraj et al., 2019
Mapping populations	317	23–154	19–121	Mahendrakar et al., 2019
	106	28–124	29–120	Kumar et al., 2016

Grain Nutrition

The goal of core breeding has been to increase yield potential in view that pearl millet has been considered as a highly nutritious cereal with higher levels of proteins and several minerals than other major cereals (Singh and Nainawatee, 1999). Earlier research reported up to 24.3% protein content in germplasm (Jambunathan and Subramanian, 1988) and up to 19.8% in elite breeding lines (Singh P. et al., 1987). However, no serious efforts were made to improve protein content because of its negative correlations with grain yield (Singh and Nainawatee, 1999). Improving grain nutritional traits is a recent addition of breeding objective, in view of global recognition of widespread deficiencies of iron (Fe) and zinc (Zn). The major areas addressed include the assessment of the extent of genetic variation for grain Fe and Zn contents, identification of diverse seed-mineral dense germplasm, nature of genotype \times environment interaction, relationships between grain minerals and agronomic traits, and genetic control of micronutrients (Govindaraj et al., 2019).

A large variability for Fe and Zn content in germplasm and breeding lines has been indicated (Table 1), suggesting the feasibility of genetic enhancement for these micronutrients. The highest Fe and Zn have been found in germplasm accessions and mapping populations derived from the *iniadi* landrace (Velu et al., 2008; Govindaraj et al., 2016). Screening of more than 120 Indian commercial hybrids has shown 46–56 ppm Fe and 37–44 ppm Zn (Rai et al., 2016). To initiate the mainstreaming of Fe and Zn in pearl millet, the levels of 42 ppm Fe and 32 ppm Zn have been set as a baseline in the Indian national testing and cultivar release policy in 2018 (AICPMIP, 2018). The daily recommended allowances for Indian adults are 17–21 and 10–12 mg/d for Fe and Zn, respectively.

Fe and Zn contents in pearl millet are largely governed by additive genetic variance (Govindaraj et al., 2013; Kanatti et al.,

2014), suggesting that both parental lines of hybrids would be required to improve for these micronutrients. Relatively lower $G \times E$ influences on the accumulation of Fe and Zn in pearl millet grains (Kanatti et al., 2014; Govindaraj et al., 2016) also indicated the effectiveness of progeny selection in the pedigree breeding to develop lines with increased levels of grain Fe and Zn densities. The higher additive genetic variance also prompts recurrent selection methods to be effective to improve the levels of grain Fe and Zn in breeding populations (Govindaraj et al., 2019).

A significant and positive association has been established between Fe and Zn (Govindaraj et al., 2013, 2016, 2020; Kanatti et al., 2014; Rai et al., 2014; Pujar et al., 2020). These two micronutrients also had a positive and highly significant correlation with seed size (Gupta et al., 2009; Kanatti et al., 2014; Govindaraj et al., 2016). Furthermore, gray and white grains from the same genetic background did not differ in their Fe/Zn levels (Govindaraj et al., 2018). These associations would give breeders leverages to develop Fe- and Zn-rich cultivars with large grain size irrespective of their color and to allow enhancement of micronutrients in mainstream breeding. The efforts in biofortification have yielded nutrient-rich and high-yield cultivars in India (Rai et al., 2014). Ten cultivars have been released with yield levels of 3.2–3.6 t/ha and 68–80 ppm Fe (Govindaraj et al., 2019), and several more are in the pipeline. Higher correlation between Fe and Zn in pearl millet exhibited that all these cultivars also had higher Zn levels (35–45 ppm Zn).

HYBRID REPLACEMENT

Replacement of old hybrids by new ones with higher potential productivity is very critical in achieving continuous genetic gains. The design followed involves the identification of potential new hybrids well in advance through multiyear and multilocal testing in the National Coordinated Trials where their performance is judged against the best hybrid released most recently (AICPMIP, 2018). The strategy essentially involves keeping a close watch on the performance and disease incidence of hybrids in their production regions. There is ~15–20 public- and 30–40 private-sector organizations engaged in meeting the national annual demand of about 22,000 metric tons seed.

Once a hybrid passes through the research and development stage of 10–15 years involving its creation, evaluation, and registration, it is introduced in the market and goes through five stages of growth, maturity, saturation, decline, and replacement. To achieve sustainable growth in the seed business, a balanced product portfolio requires a minimum of one product at each stage from introduction to decline. The hybrids are introduced in different years and are generally phased out in a 10-year time frame. An average life cycle of 6 years of top five hybrids from introduction to retirement is maintained by different public and private organizations.

Although the hybrid life cycle is largely influenced by strong competition and breakdown of resistance to DM or blast, other factors, such as market demand, alternate product options, and product quality, also play an important role. In fact, the shorter life cycle due to competition has helped to an accelerated genetic

gain due to the faster introduction of new hybrids with improved yield and resistance to diseases. This strategy has proved very critical to ensure a timely and adequate supply of desired hybrids to the farmers to sustain a continuous gain in productivity for the last few decades. With this strategy in place, ~60–70 hybrids are cultivated on the farm at any point of time in India for the last two decades.

REALIZED YIELD GAINS

Pearl millet productivity has increased from 303 kg/ha during 1950–1954 to 1,239 kg/ha during 2015–2019 that translates to an increase of more than 300% owing to the widespread use of high-yielding and disease-resistant cultivars with improved production technology. A critical analysis of the genetic improvement has been recently done in which seven decades of breeding were divided in four phases, each phase having its uniqueness (Yadav et al., 2019). During phase I (1950–1966), when genetic improvement largely concentrated on the enhancement of yield in locally adapted materials, the rate of productivity improvement was 4.5 kg/ha per year. Discovery and utilization of CMS in hybrid development marked the second phase of genetic improvement (1967–1983) in which an annual increase of 6.6 kg/ha in productivity was realized despite the large-scale cultivation of a few hybrids. A large number of genetically diverse CMS lines were developed and utilized in hybrid breeding during phase III (1984–2000), and the productivity increase was 19.0 kg/ha per year. During phase IV (2001–2018) when genetic improvement put much greater emphasis on genetic diversification of hybrids and adaptation to niche areas of cultivation, the rate of improvement in grain productivity further increased to 31.1 kg/ha per year, which is 470% of the productivity gain achieved during the first phase.

Comparison of yield increase in pearl millet vis-à-vis other four major cereals in India after the mid-1980s presents very interesting information. Following the adoption of high-yielding, disease-resistant, and stress-tolerant cultivars and crop management technology, there is a yield increase of 26% in sorghum, 59% in wheat, 69% in rice, 113% in maize, and 162% in pearl millet (**Figure 3**). These yield gains translate into 0.9% annual gains in sorghum, 2.0% in wheat, 2.3% in rice, 3.8% in maize, and 5.4% in pearl millet; all of these are much higher than the average gains achieved at the global level (FAO, 2020). The annual rate of gains in productivity is the combined outcome of improved genetics and management. This high quantum of productivity increase in pearl millet assumes greater significance in two ways. First, more than 90% of pearl millet is grown as rainfed and often on marginal lands. Second, pearl millet has attracted much lesser infrastructure and human resources in comparison to other food crops. It also affirms the correctness of priorities set in the breeding programs of India and simultaneously demonstrates the role of hybrid technology in raising crop productivity in marginal drylands.

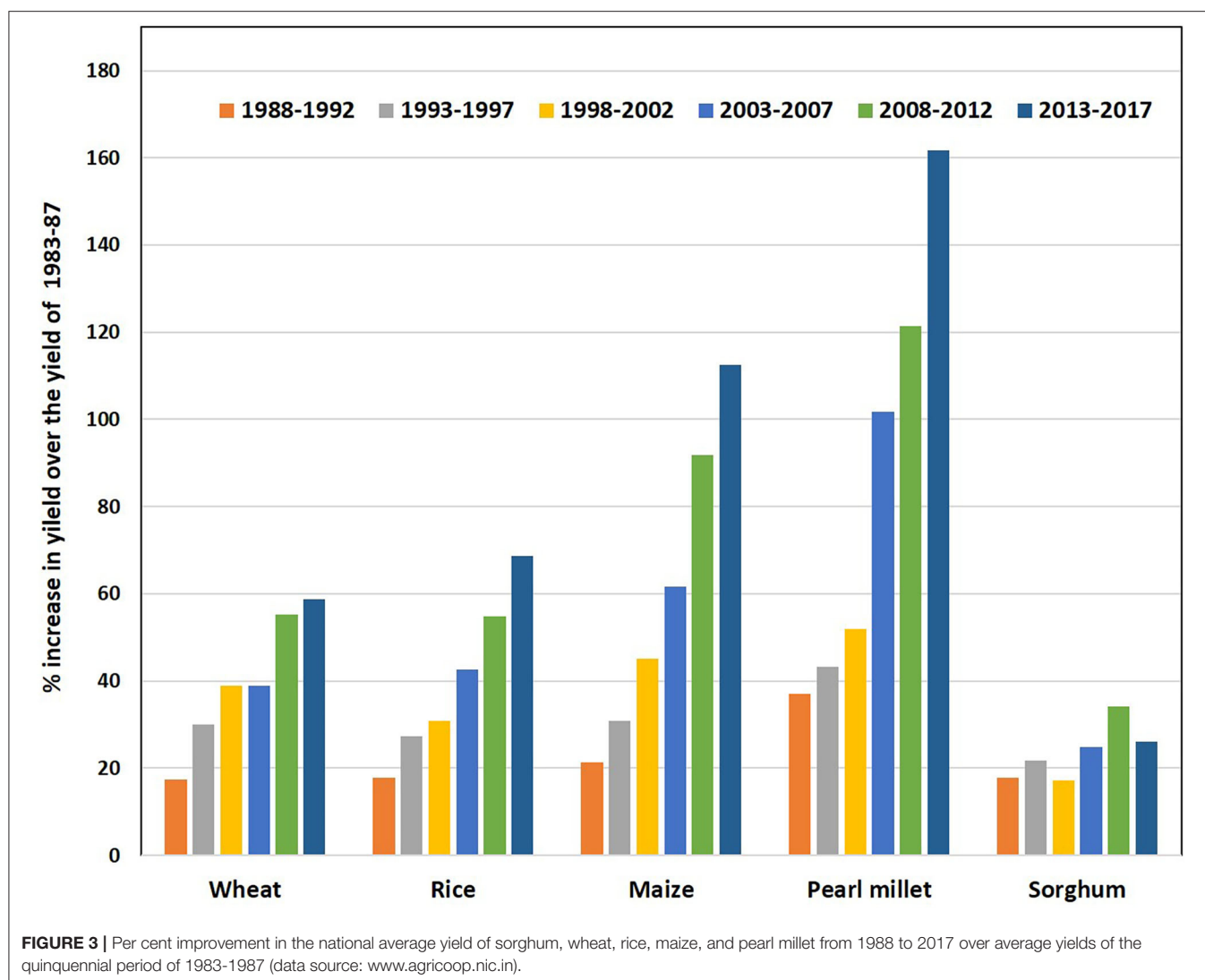
PROSPECTS OF ACCELERATING GENETIC GAINS

Pearl millet has shown impressive genetic gains in India for the past seven decades (Yadav et al., 2019). The crop now is poised to take the next quantum leap in genetic gains. The specific areas that are likely to contribute to the process are discussed in the following sections.

Genomics-Assisted Breeding

One of the most exciting developments that have implications on taking the genetic gains to the next levels in pearl millet is genomics and genomics-assisted breeding that can help improve the precision and efficiency of the breeding program. The ~1,000 genomes sequencing project has been a major milestone in pearl millet improvement (Varshney et al., 2017). This work has laid a solid foundation for carrying out trait discovery, mapping, and deployment of QTLs/alleles/candidate genes linked to traits of economic interests. It also has helped toward the development and implementation of whole-genome prediction models for the pearl millet community globally (Jarquin et al., 2020; Srivastava et al., 2020a). The whole-genome resequencing of Pearl Millet Inbred Germplasm Association Panel, mapping population parents, and elite hybrid parental lines have helped to develop a huge (>32 million) repository of genome-wide SNPs. These developments offer opportunities to rapidly map and deploy genes of agronomic importance and also to rapidly resequence lines to mine and map genes of interest. The sequencing-based mapping strategy can also help us identify superior haplotypes for different traits to form the basis of haplotype-based breeding (Sinha et al., 2020).

Many traits of agronomic importance have been mapped related to diseases, terminal drought, grain and fodder quality, combining ability loci, and heterotic gene pools (Kumar et al., 2018; Basava et al., 2019; Gupta et al., 2020; Srivastava et al., 2020a,b) with SSR and SNP markers. The available QTLs can also be remapped using the currently available SNP-based high-throughput genotyping systems. This will allow integration into the modern breeding pipelines using high-throughput genotyping platforms available currently in pearl millet (Srivastava et al., 2020a,b). Furthermore, with the availability of reference genome sequence, large-scale whole-genome resequencing data, cost-effective genotyping platform, and precise phenotyping platforms (see later), it has become possible now to map breeding-related traits in a fast manner (Bohra et al., 2020). We believe the near future will be witnessing the deployment of genomic breeding approaches such as haplotype-based breeding, forward breeding, genomic prediction, and gene editing for pearl millet improvement (Varshney et al., 2020). Along with ICRISAT, a few Indian breeding centers are using markers for the selection of terminal drought tolerance and disease resistance. However, time to release has not been significantly impacted yet, as these technologies need to be integrated with speed breeding pipelines.



Precision Phenotyping

While genotyping has become considerably cheaper and more precise in the recent past, precision phenotyping has been a major challenge, especially for drought. Full advantage of genomic resources can be taken only when quick, accurate, and cost-effective phenotypic data including root systems are available for genetic dissection of drought tolerance and selection of drought-resilient genotypes (Tuberosa, 2012). Usefulness of high-throughput and automated phenotyping platforms such as LeasyScan has been demonstrated in screening a large number of genotypes for drought tolerance (Vadez et al., 2015). There exists a much greater need to enhance the capacity for drought tolerance breeding programs to generate quick and accurate data through the use of drones, near-infrared imaging, and remote sensing.

Heterotic Grouping of Hybrid Parental Lines

Heterotic grouping of hybrid parental lines is an important strategy to increase the magnitude of heterosis on a long-term basis (Melchinger and Gumber, 1998). A diverse range of

breeding material has historically been used to develop either seed parents (B lines) or restorers parents (R lines), depending upon their specific phenotypic traits (Rai et al., 2006). Studies assessing molecular diversity classified such lines into genetically distinct groups and the confirmed existence of two broad-based groups in hybrid parents—one each for seed parents and restorer parents (Nepolean et al., 2012; Gupta et al., 2015a; Singh et al., 2018a).

Some attempts have been made in pearl millet to define heterotic groups. Pucher et al. (2016) investigated combining ability patterns in West African population hybrids but could not come out with clear heterotic groups probably because of genetic admixture in populations. Another study indicated seven heterotic groups among hybrid parents using EST and genomic SSR markers (Ramya et al., 2018). The existence of B and R lines as separate groups has been found responsible to behave as two separate broad heterotic pools, as $B \times R$ hybrids reported significantly higher levels of heterosis than $B \times B$ or $R \times R$ hybrids (Singh et al., 2018a). Recently, in a study involving 320 R and 260 B lines derived from pearl millet breeding programs in India, two each of B- and R-line heterotic groups were

identified based on the heterotic performance and combining ability (Gupta et al., 2020). Hybrids from these identified B \times R heterotic groups showed grain yield heterosis of more than 10% over the best commercial hybrid checks. This study also indicated that distinct parental groups can be formed based on molecular markers, which can help in assigning hybrid parental lines into heterotic groups to develop high-yielding hybrids. Now, the work is underway to select the appropriate testers to categorize the new hybrid parental lines or new germplasm into heterotic groups to enhance the genetic gains in pearl millet.

Combining ability studies conducted in pearl millet have shown that either there was no negative correlation or there was a positive correlation between their *per se* performance and general combining ability (GCA). Thus, high general combiners are likely to occur in lines with high grain yield *per se* than in any other yield group. Pearl millet breeding programs in India have been advancing progenies based on performance *per se* of lines during trait-based pedigree breeding, although now there is a shift toward combining ability-based selection approach. Considering the seed production economy and the high probability of producing high-yielding hybrids from high GCA lines, it is prudent that seed parents should possess both high yield *per se* and GCA.

There is not much information available about the extent of specific combining ability (SCA) and GCA variances in the existing B- and R-gene pools of pearl millet. A recent study has shown high SCA:GCA ratio (about 2 times), indicating the predominance of SCA variance over GCA variance in pearl millet hybrid parents, which was different than many of the established maize hybrid breeding programs of the United States and Europe, where low SCA:GCA ratio was observed (Gupta et al., 2020). Considering this scenario in other crops, there is a need to investigate SCA and GCA variances in the existing B- and R-line heterotic pools of pearl millet to better understand the contribution of GCA and SCA variances toward heterosis.

Harnessing Genetic Diversity

World collection of germplasm (>23,000 from 52 countries) including wild species provides a great resource to look for new sources of economic traits, disease resistance, abiotic stress tolerance, and better nutritional quality. A small fraction of germplasm has been utilized so far primarily due to the huge number of germplasm accessions and the presence of undesired traits in the unadapted genetic background. These twin problems have been largely circumvented. Development of core and minicore collections (Upadhyaya et al., 2011) is prompting breeders to use the desired germplasm in broadening the genetic base of commercial cultivars, which is very essential to reduce the chances of disease epidemics and to mitigate the effects of climate change. Very recently, >1,000 accessions of pearl millet have been sequenced as well (Varshney et al., 2017). This, in combination with genome–environment association, would help in exploring the genetic basis of adaptive traits such as drought and heat tolerance in germplasm as has been demonstrated in other crops (Frank et al., 2016; Cortés and Blair, 2018).

The accessibility of such molecular tools offers a greater opportunity to transfer specific targeted regions and to minimize

the linkage drag from agronomically inferior-looking germplasm accessions. Recently, new sources of blast resistance and flowering-stage heat tolerance have been developed in cultivated pearl millet backgrounds using wild *P. glaucum* subsp. *violaceum* (Sharma S. et al., 2020). To harness the untapped diversity, traits-specific germplasm needs to be targeted, which requires detailed characterization for new and novel traits proposed by breeders and the crop product profiles in the near future.

Addressing Host Resistance and Pathogen Virulence Together

The experience, so far, in resistance breeding for DM has indicated that most of the hybrids become susceptible in about 5–6 years of cultivation in the same area because of selection pressure in the pathogen, although there are some clear exceptions where hybrids have shown durable resistance. It would be useful to investigate resistance mechanisms operative in the parents of such hybrids to identify and deploy genes for durable DM resistance in high-yielding hybrids for the enhanced genetic gain. Continuous monitoring of virulence of pathogen populations through reaction on host differential is essential to identify resistance effective against new virulent pathotypes. Genome sequencing of pearl millet (Varshney et al., 2017) and pathogens of DM and blast (Nayaka et al., 2017; Prakash et al., 2019) will help in understanding the molecular basis of compatible/incompatible host \times pathogen interaction and provide a greater opportunity to breeders and pathologists to control the diseases. With the availability of genomic tools, identification of QTLs determining resistance to particular disease, and demonstration of the success of marker-assisted backcrossing, it now appears possible to stack target QTLs in the parental lines of commercial hybrids having multiple resistance to various pathotypes of DM, blast, and rust to realize a greater genetic gain. It should be recognized that QTLs transferred through backcrossing are pathotype-specific that indicates vertical resistance of their nature. These impart relatively large phenotypic variance and hence may house major genes for disease resistance. These may, however, get defeated by the pathogen relatively quickly; hence, identification of QTLs responsible for horizontal resistance is underway.

Strengthening Hybrid Breeding for Arid Regions

One of the key issues, often debated in past, has been the comparative advantage of hybrids or Open-pollinated varieties (OPVs) under severe drought conditions, given the reports that genetically heterogeneous OPVs might exploit population buffering mechanism (Bradshaw, 1965; Haussmann et al., 2000) to provide stable performance under unpredictable drought environments. A comprehensive study conducted in arid regions of India comparing 142 hybrids and 84 composites over 12 years in 94 environments for their performance reported that hybrids yielded significantly higher grain than composites with an overall superiority of 25% (Yadav et al., 2012a).

Other consideration, in cultivar adoption in the existing seed supply system of pearl millet, is the multiple sowings, especially

in drought areas. Composites have an edge over hybrids as they are self-perpetuating, and harvested seed can be used to plant the next crop. However, such an option would come with a significant penalty at least for grain yield. Therefore, hybrids are likely to play a much greater role than composites in enhancing pearl millet productivity in drought-prone regions including SSA.

Mainstreaming the Biofortification

Mainstreaming of Fe and Zn in pearl millet breeding needs to be implemented to achieve nutritional security in SA and SSA. Breeding for micronutrients and vitamins has been initiated by HarvestPlus, a CGIAR Challenge Program. In collaborations with national partners, ICRISAT has generated now enough database for Fe and Zn. Greater than 60 ppm for Fe and 40 ppm for Zn were targeted for breeding as a part of the ICRISAT product profile. This is a momentous step toward mainstreaming in elite breeding lines. The dissemination of these mainstreamed breeding lines and hybrid parents and their utilization (both public and private sector) would make biofortified hybrid development a regular activity to enhance genetic gains for the micronutrient traits in the long term.

This is very encouraging that no negative association has been reported between grain yield and micronutrients in pearl millet, suggesting the feasibility of combining high yield with a greater concentration of micronutrients. Efficacy studies have indicated bioavailability of 7% of Fe and 25% of Zn in pearl millet. Biofortified pearl millet may provide up to 80% of Fe and 100% of Zn daily requirements and would have a greater impact on human health.

Improving Nutrient Use Efficiency

Although pearl millet is mainly cultivated on sandy and sandy-loam soils that are inherently low in their nitrogen (N) and phosphorous (P) contents, its adaptation to low nutrients is seldom addressed assuming that this issue can be easily addressed through the external use of fertilizers. Limited studies conducted on this aspect (Gemenet et al., 2015) have indicated the possibility of breeding nutrient-use-efficient (NUE) cultivars of pearl millet. Looking to soil degradation and water contamination due to the leaching of N in subsurface or groundwater, pearl millet can be an important source of native genes that confer adaptation to low nutrient conditions (Serba et al., 2020). There is a need for a systematic study to understand the relevant traits' priority and their magnitude of variability for NUE using core breeding materials including minicore collections available at GenBank (Pujarula et al., 2021).

Synergizing Breeding and Agronomics

A sustained increase in pearl millet productivity requires the integration of suitable cultural practices in its diverse production environments for disease resistant and improved cultivars to achieve greater genetic gains. On-farm demonstrations of improved cultivars and production technologies have established that the pearl millet yields at farm levels can easily be enhanced by 20–25% by adopting suitable agrotechniques (Yadav et al., 2012a). Intensive management including higher planting

density, irrigation scheduling, and recommended use of mineral fertilizers in better-endowed areas would play a critical role in harnessing the potential yield of improved cultivars. Widely spaced crop, integration of legumes in pearl millet-based cropping system to maintain soil fertility, and microdosing of nutrients are very important to further enhance productivity gains in drought-prone regions. Machine harvestable plant type and lodging resistance are the need of the hour in reducing cultivation cost and enhancing profitability.

Speed Breeding and Big-Data Analytics

Genetic gains of any breeding program significantly depend upon the number of crop breeding cycles a program can undertake in a year. This varies in different breeding programs as per their local weather conditions. For instance, only one crop of pearl millet can be taken in north India, while two crops per year can be grown in western, central, and peninsular India. Under this current scenario, breeding a new crop cultivar takes about a decade or more, with 6 or 7 years spent in seasonal generational advancements to arrive at elite materials that go for testing and release. Now, new environmentally controlled facilities, known as “RapidGen,” have been developed, which will shorten the 6–7-year window significantly. When used with the full suite of breeding acceleration techniques, RapidGen can make it possible to take four crop cycles in a year (<https://www.icrisat.org/first-public-research-facility-to-put-agriculture-on-fast-forward-launched-at-icrisat/>). With such new facilities in place, we are now moving toward the new era of speed breeding in pearl millet where genetic gains are poised to take a further leap.

Over the past several decades, judicious use of data analytics in multilocal trials and quantitative genetics played a major role in achieving higher genetic gains in pearl millet. However, the present era is of ultrahigh-speed computing, crop simulations, big-data analytics, internet of things, artificial intelligence, and machine learning and must be exploited in pearl millet breeding for achieving better genetic gains. These high-throughput streams need to be decoded for pearl millet improvement by developing an appropriate digital data capture platform, breeding databases, modern quantitative genetics, and real-time analytics. All such information in databases will provide an opportunity to run complex queries and scenario analysis enabling researchers to focus on specialized research. One extremely computationally intensive data science intervention will be the use of quantitative genetics-based crop simulation algorithms to understand and optimize existing pearl millet breeding pipelines and take measures to refine them further.

Building Partnership

Pearl millet research and development are a mandate of several national and international organizations with a common goal of making an impact on the communities cultivating pearl millet. With this shared goal, the partnership needs to be pursued systematically. Partnership with international and advanced research institutes has contributed to providing access to well-characterized genetic resources and developing genomic resources. Partnership with the private sector has been most

critical in delivering the products of improved genetics to the farming community. One such successful existing example of public-private partnership (PPP) is the ICRISAT-Private Sector Pearl Millet Seed Companies Consortium, established in 2000 and which has an engagement of ~30 seed companies. This consortium has proved to be a very effective vehicle to quickly deliver research outputs of ICRISAT's breeding program (Gowda et al., 2009). Another successful example of PPP is the International Pearl millet Genome Consortium, where 65 scientists from 30 research institutions across the world came together and unraveled the sequence of the pearl millet genome (Varshney et al., 2017). Such partnerships hold the key in providing adequate resources and making advances in cutting-edge science technologies to realize a sustained growth in pearl millet productivity.

CONCLUSION

Pearl millet is becoming an indispensable food crop that provides calories, nutrition, and livelihood security to the poor and marginal people living in the fragile ecosystem of arid and

semiarid regions of SA and SSA. Pearl millet is a crop of choice because of its critical role in enhancing the resilience to climate change. The past breeding priorities and strategies have been able to deliver significant productivity growth realized in pearl millet. Greater use of hybrid technology, employing modern tools, wider interinstitutional, and intersectoral partnerships, and improved crop management practices would play a greater role to accomplish much higher genetic gains for yield and nutritional traits for growing populations in SA and SSA. The success would depend upon a deeper understanding of new germplasm, genome, and trait-specific genes for novel traits through an amalgamation of conventional and modern tools and rapid generation techniques in national and international pearl millet breeding programs.

AUTHOR CONTRIBUTIONS

OY conceived the idea of writing this review and prepared the draft of the manuscript. SG, MG, RS, RV, RKS, AR, and RM contributed and strengthened different sections included in the review. All authors read and approved the manuscript.

REFERENCES

- Abraham, P., Alimpta, P. D., and Bdliya, B. S. (2019). Inheritance of resistance to ergot disease in a diallel cross of pearl millet (*Pennisetum glaucum* (L.) R. Br.). *Tanzania J. Agric. Sci.* 18, 50–58.
- AICPMIP (2018). *Proceedings of the 53rd Annual Group Meeting of ICAR - All India Coordinated Research Project on Pearl Millet (AICPMIP)*. Jodhpur: Agriculture University. Available online at: <http://www.aicpmip.res.in/pw2018.pdf> (accessed July 7, 2019)
- Andrews, D. J., and Anand Kumar, K. (1996). Use of the West African pearl millet landrace Iniadi in cultivar development. *Plant Genetic Resour. Newsl.* 105, 15–22.
- Andrews, D. J., and Kumar, K. A. (1992). Pearl millet for food, feed and forage. *Adv. Agron.* 48, 89–139. doi: 10.1016/S0065-2113(08)60936-0
- Athwal, D. S. (1961). Recent developments in the breeding and improvement of *bajra* (pearl millet) in the Punjab. *Madras Agric. J.* 48, 18–19.
- Athwal, D. S. (1966). Current plant breeding research with special reference to *Pennisetum*. *Indian J. Genet. Plant Breed.* 26A, 73–85.
- Basava, R. K., Hash, C. T., Mahendrakar, M. D., Kishor, P. B., K., Satyavathi, C. T., et al. (2019). Discerning combining ability loci for divergent environments using chromosome segment substitution lines (CSSLs) in pearl millet. *PLoS ONE* 14:e0218916. doi: 10.1371/journal.pone.0218916
- Bidinger, F. R., Mahalakshmi, V., and Rao, G. D. P. (1987). Assessment of drought resistance in pearl millet [*Pennisetum americanum* (L.) Leeke]: I. Factors affecting yield under stress. *Aust. Agric. Res.* 38, 37–48. doi: 10.1071/AR9870037
- Bidinger, F. R., Nepolean, T., Hash, C. T., Yadav, R. S., and Howarth, C. J. (2007). Quantitative trait loci for grain yield in pearl millet under variable post-flowering moisture conditions. *Crop Sci.* 47, 969–980. doi: 10.2135/cropsci2006.07.0465
- Bohra, A., Jha, U. C., Godwin, I. D., and Varshney, R. K. (2020). Genomic interventions for sustainable agriculture. *Plant Biotech. J.* 18:12. doi: 10.1111/pbi.13472
- Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Adv. Genet.* 13, 115–155. doi: 10.1016/S0065-2660(08)60048-6
- Burton, G. W. (1965). Pearl millet Tift 23A release. *Crops Soils* 17:19.
- Cortés, A. J., and Blair, M. W. (2018). Genotyping by sequencing and genome-environment associations in wild common bean predict widespread divergent adaptation to drought. *Front. Plant Sci.* 9:128. doi: 10.3389/fpls.2018.00128
- Djanaguiraman, M., Perumal, R., Ciampitti, I. A., Gupta, S. K., and Prasad, P. V. V. (2018). Quantifying pearl millet response to high temperature stress: thresholds, sensitive stages, genetic variability and relative sensitivity of pollen and pistil. *Plant Cell Environ.* 41, 993–1007. doi: 10.1111/pce.12931
- FAO (2020). Available online at: <http://www.fao.org/faostat/en/#data/QC> (accessed July 30, 2020).
- Frank, A., Oddou-Muratorio, S., Lalagüe, H., Pluess, A. R., Heiri, C., and Vendramin, G. G. (2016). Genome-environment association study suggests local adaptation to climate at the regional scale in *Fagus sylvatica*. *New Phytol.* 210, 589–601. doi: 10.1111/nph.13809
- Fussell, L. K., Bidinger, F. R., and Bieler, P. (1991). Crop physiology and breeding for drought tolerance. Research and development. *Field Crops Res.* 27, 183–199. doi: 10.1016/0378-4290(91)90061-Y
- Gemenet, D. C., Hash, C. T., Sanogo, M. D., Sy, O., Zangre, R. G., Leiser, W. L., et al. (2015). Phosphorus uptake and utilization efficiency in West African pearl millet inbred lines. *Field Crops Res.* 171, 54–66. doi: 10.1016/j.fcr.2014.11.001
- Gill, K. S. (1991). *Pearl Millet and its Improvement*. New Delhi: Indian Council of Agricultural Research, 305.
- Govindaraj, M., Kanatti, A., Rai, K. N., and Satyanarayana, T. (2019). Does soil micronutrient variability in test locations influence performance of biofortified pearl millet in India? *Better Crops South Asia* 11, 31–35. Available online at: <http://oar.icrisat.org/11285>
- Govindaraj, M., Rai, K. N., and Shanmugasundaram, P. (2016). Intra-population genetic variance for grain iron and zinc contents and agronomic traits in pearl millet. *Crop J.* 4, 48–54. doi: 10.1016/j.cj.2015.11.002
- Govindaraj, M., Rai, K. N., Shanmugasundaram, P., Dwivedi, S. L., Sahrawat, K. L., Muthaiah, A. R., et al. (2013). Combining ability and heterosis for grain iron and zinc densities in pearl millet. *Crop Sci.* 53, 505–517. doi: 10.2135/cropsci2012.08.0477
- Govindaraj, M., Rao, A. S., Shivade, H., and Rai, K. N. (2018). Effect of grain colour on iron and zinc density in pearl millet. *Indian J. Genet. Plant Breed.* 78, 247–251. doi: 10.5958/0975-6906.2018.00032.9
- Govindaraj, M., Virk, P. S., Kannati, A., Cherian, B., Rai, K. N., Anderson, M. S., et al. (2020). "Biofortified pearl millet cultivars offer potential solution to tackle malnutrition in India," in *Quantitative Genetics, Genomics and Plant Breeding*, ed Manjit S. Kang (Oxon: CAB International), 385–396. doi: 10.1079/9781789240214.0385

- Gowda, C. L. L., Verma, P., Rai, K. N., Reddy, B. V. S., and Saxena, K. B. (2009). "Private sector partnering on crops for the poorest of the poor," in *World Food Security: Can Private Sector R and D Feed the Poor?* (Canberra).
- Gupta, S. K., Ameta, V. L., Pareek, S., Mahala, R. S., Jayalekha, A. K., Deora, V. S., et al. (2016). "Genetic enhancement for flowering period heat tolerance in pearl millet (*Pennisetum glaucum* L. (R.) Br.)," in *7th International Crop Science Congress* (Beijing). Available online at: <http://oar.icrisat.org/9757/1/Page1.pdf> (accessed December 12, 2018)
- Gupta, S. K., Ameta, V. L., Pareek, S., Singh, I., Deora, V. S., Verma, Y. S., et al. (2019). "Enhancing flowering period heat tolerance in pearl millet through shuttle breeding," in *Presented at 3rd Agriculture and Climate Change Conference* (Budapest). Available online at: https://www.elsevier.com/_data/assets/pdf_file/0010/847882/AGRI-2019-Oral-Programme_v5_DE.pdf (accessed January 02, 2020)
- Gupta, S. K., Nepolean, T. V., Sankar, S. M., Rathore, A., Das, R. R., Rai, K. N., et al. (2015a). Patterns of molecular diversity in current and previously developed hybrid parents of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Am. J. Plant Sci.* 06, 1697–1712. doi: 10.4236/ajps.2015.611169
- Gupta, S. K., Patil, K. S., Rathore, A., Yadav, D. V., Sharma, L. D., Mungra, K. D., et al. (2020). Identification of heterotic groups in South-Asian-bred hybrid parents of pearl millet. *Theor. Appl. Genet.* 133, 873–888. doi: 10.1007/s00122-019-03512-z
- Gupta, S. K., Rai, K. N., Govindaraj, M., and Rao, A. S. (2012a). Genetics of fertility restoration of the A4 cytoplasmic-nuclear male sterility system in pearl millet. *Czech J. Genet. Plant Breed.* 48, 87–92. doi: 10.17221/164/2011-CJGPB
- Gupta, S. K., Rai, K. N., Singh, P., Ameta, V. L., Gupta, S. K., Jayalekha, A. K., et al. (2015b). Seed set variability under high temperatures during flowering period in pearl millet (*Pennisetum glaucum* L. (R.), Br.). *Field Crops Res.* 171, 41–53. doi: 10.1016/j.fcr.2014.11.005
- Gupta, S. K., Rathore, A., Yadav, O. P., Rai, K. N., Khairwal, I. S., Rajpurohit, B. S., et al. (2013). Identifying mega-environments and essential test locations for pearl millet cultivar selection in India. *Crop Sci.* 53, 2444–2453. doi: 10.2135/cropsci2013.01.0053
- Gupta, S. K., Sharma, R., Rai, K. N., and Thakur, R. P. (2012b). Inheritance of foliar blast resistance in pearl millet (*Pennisetum glaucum*). *Plant Breed.* 131, 217–219. doi: 10.1111/j.1439-0523.2011.01929.x
- Gupta, S. K., Velu, G., Rai, K. N., and Sumalini, K. (2009). Association of grain iron and zinc content with grain yield and other traits in pearl millet (*Pennisetum glaucum* (L.) R. Br.). *Crop Improv.* 6, 4–7. Available online at: <http://oar.icrisat.org/id/eprint/363>
- Gupta, S. K., Yadav, D. V., Govindaraj, M., Boratkar, M., Kulkarni, V. N., and Rai, K. N. (2018). Inheritance of fertility restoration of A5 cytoplasmic-nuclear male sterility system in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Indian J. Genet. Plant Breed.* 78, 228–232. doi: 10.5958/0975-6906.2018.00029.9
- Hanna, W. W. (1989). Characteristics and stability of a new cytoplasmic-nuclear male sterile source in pearl millet. *Crop Sci.* 29, 1457–1459. doi: 10.2135/cropsci1989.0011183X002900060026x
- Hash, C. T., Sharma, A., Kolesnikova-Allen, M. A., Singh, S. D., Thakur, R. P., Raj, A. G. B. (2006). Teamwork delivers biotechnology products to Indian small-holder crop-livestock producers: pearl millet hybrid "HHB 67 Improved" enters seed delivery pipeline. *J. SAT Agric. Res.* 2, 16–20. Available online at: <http://oar.icrisat.org/id/eprint/2738>
- Hausmann, B. I. G., Obilana, A. B., Ayiecho, P. O., Blum, A., Schipprack, W., and Geiger, H. H. (2000). Yield and yield stability of four population types of grain sorghum in a semi-arid area of Kenya. *Crop Sci.* 40, 319–329. doi: 10.2135/cropsci2000.402319x
- Howarth, C. J., Pollock, C. J., and Peacock, J. M. (1997). Development of laboratory-based methods for assessing seedling thermotolerance in pearl millet. *New Phytol.* 137, 129–139. doi: 10.1046/j.1469-8137.1997.00827.x
- Jambunathan, R., and Subramanian, V. (1988). "Grain quality and utilization of sorghum and pearl millet," in *Biotechnology in Tropical Crop Improvement*, eds J. M. J. de Wet and T. A. Preston (Patancheru: ICRISAT), 133–139.
- Jarquín, D., Howard, R., Liang, Z., Gupta, S. K., Schnable, J. C., and Crossa, J. (2020). Enhancing hybrid prediction in pearl millet using genomic and/or multi-environment phenotypic information of inbreds. *Front. Genet.* 10:1294. doi: 10.3389/fgene.2019.01294
- Jukanti, A., Gowda, C. L. L., Rai, K. N., Manga, V. K., and Bhatt, R. K. (2016). Crops that feed the world 11. Pearl Millet (*Pennisetum glaucum* L.): an important source of food security, nutrition and health in the arid and semi-arid tropics. *Food Sec.* 8, 307–329. doi: 10.1007/s12571-016-0557-y
- Kanatti, A., Rai, K. N., Radhika, K., Govindaraj, M., Sahrawat, K. L., and Rao, A. S. (2014). Grain iron and zinc density in pearl millet: combining ability, heterosis and association with grain yield and grain size. *Springerplus* 3, 1–12. doi: 10.1186/2193-1801-3-763
- Knox, J. W., Hess, T. M., Daccache, A., and Perez Ortola, M. (2011). *What are the projected impacts of climate change on food crop productivity in Africa and South Asia?* DFID Systematic Review Final Report, 77. Cranfield University, Cranfield, United States.
- Kumar, S., Hash, C., Nepolean, T., Mahendrakar, M., Satyavathi, C., Singh, G., et al. (2018). Mapping grain iron and zinc content quantitative trait loci in an inbred-derived immortal population of pearl millet. *Genes (Basel)* 9:248. doi: 10.3390/genes9050248
- Kumar, S., Hash, C. T., Thirunavukkarasu, N., Singh, G., Rajaram, V., Rathore, A., et al. (2016). Mapping quantitative trait loci controlling high iron and zinc in self and open pollinated grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 7:1636. doi: 10.3389/fpls.2016.01636
- Lynch, P. J. (1994). Recurrent selection in pearl millet for improving stand establishment at high temperature. *Plant Breed.* 113, 237–241. doi: 10.1111/j.1439-0523.1994.tb00728.x
- Mahendrakar, M. D., Kumar, S., Singh, R. B., Rathore, A., Potupureddi, G., Kavi Kishor, P. B., et al. (2019). Genetic variability, genotype × environment interaction and correlation analysis for grain iron and zinc contents in recombinant inbred line population of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Indian J. Genet.* 79, 545–551. doi: 10.31742/IJGPB.79.3.3
- Melchinger, A. E., and Gumber, R. K. (1998). "Overview of heterosis and heterotic groups in agronomic crops," in *Concepts and Breeding of Heterosis in Crop Plants*, eds K. R. Lamkey and J. E. Staub (Madison: CSSA), 29–44. doi: 10.2135/cssaspecpub25.c3
- Nayaka, S. C., Shetty, H. S., Satyavathi, C. T., Yadav, R. S., Kishor, P. K., Nagaraju, M., et al. (2017). Draft genome sequence of *Sclerospora graminicola*, the pearl millet downy mildew pathogen. *Biotechnol. Rep.* 16, 18–20. doi: 10.1016/j.btre.2017.07.006
- Nepolean, T., Gupta, S. K., Dwivedi, S. L., Bhattacharjee, R., Rai, K. N., and Hash, C. T. (2012). Genetic diversity in maintainer and restorer lines of pearl millet. *Crop Sci.* 52, 2555–2563. doi: 10.2135/cropsci2011.11.0597
- Patil, K. S., Gupta, S. K., Marathi, B., Danam, S., Thatikunta, R., Rathore, A., et al. (2020). African and Asian origin pearl millet populations: genetic diversity pattern and its association with yield heterosis. *Crop Sci.* 60, 3035–3048. doi: 10.1002/csc2.20245
- Peacock, J. M., Soman, P., Jayachandran, R., Rani, A. U., Howarth, C. J., and Thomas, A. (1993). Effects of high soil surface temperature on seedling survival in pearl millet. *Exp. Agric.* 29, 215–225. doi: 10.1017/S0014479700020664
- Prakash, G., Kumar, A., Sheoran, N., Aggarwal, R., Satyavathi, C. T., Chikara, S. K., et al. (2019). First draft genome sequence of a pearl millet blast pathogen, *Magnaporthe grisea* strain PMg_DL, obtained using PacBio single-molecule real-time and illumina NextSeq 500 sequencing. *Microb. Resour. Announc.* 8:e01499-18. doi: 10.1128/MRA.01499-18
- Presterl, T., and Weltzien, E. (2003). Exploiting heterosis in pearl millet population breeding in arid environments. *Crop Sci.* 43, 767–776. doi: 10.2135/cropsci2003.7670
- Pucher, A., Sy, O., Sanogo, M. D., Angarawai, I. I., Zangre, R., Ouedraogo, M., et al. (2016). Combining ability patterns among West African pearl millet landraces and prospects for pearl millet hybrid breeding. *Field Crop Res.* 195, 9–20. doi: 10.1016/j.fcr.2016.04.035
- Pujar, M., Govindaraj, M., Gangaprasad, S., Kanatti, A., and Shivade, H. (2020). Genetic variation and diversity for grain iron, zinc, protein and agronomic traits in advanced breeding lines of pearl millet [*Pennisetum glaucum* (L.) R. Br.] for biofortification breeding. *Genetic Resour. Crop Evol.* 67, 2009–2022. doi: 10.1007/s10722-020-00956-x
- Pujarula, V., Pusuluri, M., Bolla, S., Das, R. R., Ratnala, R., Adapala, G., et al. (2021). Genetic variation for nitrogen use efficiency traits in global diversity panel and parents of mapping populations in pearl millet. *Front. Plant Sci.* 12:625915. doi: 10.3389/fpls.2021.625915
- Raghvani, K. L., Juneja, R. P., Ghelani, Y. H., Parmar, G. M., and Dangaria, C. J. (2008). Influence of abiotic factors on population fluctuations of major insect pest of pearl millet. *Indian J. Appl. Entom.* 22, 48–50.

- Rai, K. N. (1995). A new cytoplasmic-nuclear male sterility system in pearl millet. *Plant Breed.* 114, 445–447. doi: 10.1111/j.1439-0523.1995.tb00829.x
- Rai, K. N., and Anand Kumar, K. (1994). Pearl millet improvement at ICRISAT - an update. *Intern. Sorghum Millets Newsl.* 35, 1–29.
- Rai, K. N., Gowda, C. L. L., Reddy, B. V. S., and Sehgal, S. (2008). The potential of sorghum and pearl millet in alternative and health food uses. *Comp. Rev. Food Sci. Food Saf.* 7, 340–352. doi: 10.1111/j.1541-4337.2008.00049.x
- Rai, K. N., Gupta, S. K., Bhattachajee, R., Kulkarni, V. N., Singh, A. K., and Rao, A. S. (2009a). "Morphological characteristics of ICRISAT-bred pearl millet hybrid parents," in *International Crops Research for the Semi-arid Tropics* (Patancheru), 176.
- Rai, K. N., Gupta, S. K., Sharma, R., Govindaraj, M., Rao, A. S., Shivade, H., et al. (2014). Pearl millet breeding lines developed at ICRISAT: a reservoir of variability and useful source of non-target traits. *SAT eJournal* 1, 1–13. Available online at: <http://oar.icrisat.org/id/eprint/8629>
- Rai, K. N., and Hanna, W. W. (1990). Morphological characteristics of tall and dwarf pearl millet isolines. *Crop Sci.* 30, 23–25. doi: 10.2135/cropsci1990.0011183X003000010005x
- Rai, K. N., Khairwal, I. S., Dangaria, C. J., Singh, A. K., and Rao, A. S. (2009b). Seed parent breeding efficiency of three diverse cytoplasmic-nuclear male sterility systems in pearl millet. *Euphytica* 165, 495–507. doi: 10.1007/s10681-008-9765-7
- Rai, K. N., Kulkarni, V. N., Thakur, R. P., Haussmann, B. I. G., and Mgonja, M. A. (2006). "Pearl millet hybrid parent's research: approaches and achievements," in *Hybrid Parents Research at ICRISAT*, eds C. L. L. Gowda, K. N. Rai, B. V. S. Reddy, and K. B. Saxena (Patancheru: International Crops Research Institute for the Semi-Arid Tropics), 11–74.
- Rai, K. N., and Rao, A. S. (1991). Effect of d2 dwarfing gene on grain yield and yield components in pearl millet near-isogenic lines. *Euphytica* 52, 25–31.
- Rai, K. N., and Thakur, R. P. (1996). Smut reaction of pearl millet hybrids affected by fertility restoration and genetic resistance of parental lines. *Euphytica* 90, 31–37.
- Rai, K. N., Virk, D. S., Harinarayana, G., and Rao, A. S. (1996). Stability of male sterile sources and fertility restoration of their hybrids in pearl millet. *Plant Breed.* 115, 494–500. doi: 10.1111/j.1439-0523.1996.tb00964.x
- Rai, K. N., Yadav, O. P., Govindaraj, M., Pfeiffer, W. H., Yadav, H. P., Rajpurohit, B. S., et al. (2016). Grain iron and zinc densities in released and commercial cultivars of pearl millet (*Pennisetum glaucum*). *Indian J. Agric. Sci.* 86, 291–296.
- Rai, K. N., Yadav, O. P., Gupta, S. K., Mahala, R. S., and Gupta, S. K. (2012). Emerging research priorities in pearl millet. *J. SAT Agric. Res.* 10, 1–4. Available online at: <http://ejournal.icrisat.org/index.htm>
- Rama Rao, C. A., Raju, B. M. K., Rao, A. V. M. S., Reddy, D. Y., Meghana, Y. L., Swapna, N., et al. (2019). Yield variability of sorghum and pearl millet to climate change in India. *Indian J. Agril. Econ.* 74, 350–362.
- Ramya, A. R., Ahamed, M. L., Satyavathi, C. T., Rathore, A., Katiyar, P., Raj, A. G., et al. (2018). Towards defining heterotic gene pools in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 8:1934. doi: 10.3389/fpls.2017.01934
- Rattunde, H. F., Singh, P., and Witcombe, J. R. (1989). Feasibility of mass selection in pearl millet. *Crop Sci.* 29, 1423–1427. doi: 10.2135/cropsci1989.0011183X002900060018x
- Safeulla, K. M. (1976). *Biology and Control of the Downy Mildews of Pearl Millet, Sorghum and Pearl Millet*. Mysore: Wesley Press, 304.
- Serba, D. D., and Yadav, R. S. (2016). Genomic tools in pearl millet breeding for drought tolerance: status and Prospects. *Front. Plant Sci.* 7:1724. doi: 10.3389/fpls.2016.01724
- Serba, D. D., Yadav, R. S., Varshney, R. K., Gupta, S. K., Mahalingam, G., Srivastava, R. K., et al. (2020). "Genomic designing of pearl millet: a resilient crop for arid and semi-arid environments," in *Genomic Designing of Climate-Smart Cereal Crops*, ed C. Kole (Cham: Springer), 221–286. doi: 10.1007/978-3-319-93381-8_6
- Sharma, H. C., and Youm, O. (1999). "Host plant resistance in integrated pest management," in *Pearl Millet Breeding*, eds I. S. Khairwal, K. N. Rai, D. J. Andrews, and G. Harinarayana (New Delhi: Oxford and IBH Co, Pvt. Limited), 381–418.
- Sharma, P. C., Singh, D., Sehgal, D., Singh, G., Hash, C. T., and Yadav, R. S. (2014). Further evidence that a terminal drought tolerance QTL of pearl millet is associated with reduced salt uptake. *Environ. Exp. Bot.* 102, 48–57. doi: 10.1016/j.envexpbot.2014.01.013
- Sharma, R., Sharma, S., and Gate, V. L. (2020). Tapping *Pennisetum violaceum*, a wild relative of pearl millet (*Pennisetum glaucum*), for resistance to blast (caused by *Magnaporthe grisea*) and rust (caused by *Puccinia substriata* var. indica). *Plant Dis.* 104, 1487–1491. doi: 10.1094/PDIS-08-19-1602-RE
- Sharma, R., Upadhyaya, H. D., Manjunatha, S. V., Rai, K. N., Gupta, S. K., and Thakur, R. P. (2013). Pathogenic variation in the pearl millet blast pathogen, *Magnaporthe grisea* and identification of resistance to diverse pathotypes. *Plant Dis.* 97, 89–95. doi: 10.1094/PDIS-05-12-0481-RE
- Sharma, S., Sharma, R., Pujar, M., Yadav, D., Yadav, Y. P., Rathore, A., et al. (2020). Utilization of wild *Pennisetum* species for improving biotic and abiotic stress tolerance in pearl millet (*Pennisetum glaucum* L.). *Crop Sci.* 61, 289–304. doi: 10.1002/csc2.20408
- Singh, F., and Nainawatee, H. S. (1999). "Grain quality traits," in *Pearl Millet Breeding*, eds I. S. Khairwal, K. N. Rai, D. J. Andrews, and G. Harinarayana (New Delhi: Oxford and IBH), 157–183.
- Singh, P., Singh, U., Eggum, B. O., Anand Kumar, K., and Andrews, D. J. (1987). Nutritional evaluation of high protein genotypes of pearl millet (*Pennisetum americanum* (L.) Leeke). *J. Sci. Food Agric.* 38, 41–48. doi: 10.1002/jsfa.2740380108
- Singh, S., Gupta, S. K., Thudi, M., Das, R. R., Vemula, A., Garg, V., et al. (2018a). Genetic diversity patterns and heterosis prediction based on SSRs and SNPs in hybrid parents of pearl millet. *Crop Sci.* 58, 2379–2390. doi: 10.2135/cropsci2018.03.0163
- Singh, S., Sharma, R., Pushpavathi, B., Gupta, S. K., Durgarani, C. V., and Raj, C. (2018b). Inheritance and allelic relationship among gene(s) for blast resistance in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Plant Breed.* 137, 573–584. doi: 10.1111/pbr.12607
- Singh, S. D. (1995). Downy mildew of pearl millet. *Plant Dis.* 79, 545–550. doi: 10.1094/PD-79-0545
- Singh, S. D., Andrews, D. J., and Rai, K. N. (1987). Registration of ICML 11 rust-resistant pearl millet germplasm. *Crop Sci.* 27, 367–368. doi: 10.2135/cropsci1987.0011183X002700020062x
- Singh, S. D., King, S. B., and Malla Reddy, P. (1990). Registration of five pearl millet germplasm sources with stable resistance to rust. *Crop Sci.* 30:1165. doi: 10.2135/cropsci1990.0011183X003000050061x
- Singh, S. D., Wilson, J. P., Navi, S. S., Talukdar, B. S., Hess, D. E., and Reddy, K. N. (1997). "Screening techniques and sources of resistance to downy mildew and rust in pearl millet," in *Information Bulletin no. 48. International Crops Research Institute for the Semi-Arid Tropics* (Patancheru), 36.
- Sinha, P., Singh, V. K., Saxena, R. K., Khan, A. W., Abbai, R., Chitikineni, A., et al. (2020). Superior haplotypes for haplotype-based breeding for drought tolerance in pigeonpea (*Cajanus cajan* L.). *Plant Biotech. J.* 1:1. doi: 10.1111/pbi.13422
- Soman, P., Biding, F. R., Peacock, J. M., and Walker, T. S. (1981). *Seedling Establishment - A Preliminary Survey Taken Up in Aurepally During kharif 1981*. ICRISAT Internal Report. ICRISAT, Patancheru, India.
- Soman, P., Jayachandran, R., and Biding, F. R. (1987). Uneven variation in plant to plant spacing in pearl millet. *Agron. J.* 79, 891–895. doi: 10.2134/agronj1987.00021962007900050027x
- Soman, P., and Peacock, J. M. (1985). A laboratory technique to screen seedling emergence of sorghum and pearl millet at high soil temperature. *Exp. Agric.* 21, 335–341. doi: 10.1017/S0014479700013168
- Srivastava, R. K., Bollam, S., Pujarula, V., Pusuluri, M., Singh, R. B., Potupureddi, G., et al. (2020b). Exploitation of heterosis in pearl millet: a review. *Plants* 9:807. doi: 10.3390/plants9070807
- Srivastava, R. K., Singh, R. B., Srikanth, B., Satyavathi, C. T., Yadav, R., and Gupta, R. (2020a). Genome-wide association studies (GWAS) and genomic selection (GS) in pearl millet: advances and prospects. *Front. Genet.* 10:1389. doi: 10.3389/fgene.2019.01389
- Sultan, B., Roudier, P., Quirion, P., Alhassane, A., Muller, B., Dingkuhn, M., et al. (2013). Assessing climate change impacts on sorghum and millet yields in the Sudanian and Sahelian savannas of West Africa. *Environ. Res. Lett.* 8:014040. doi: 10.1088/1748-9326/8/1/014040
- Thakur, R. P., King, S. B., Rai, K. N., and Rao, V. P. (1992). "Identification and utilization of smut resistance in pearl millet," in *Research Bulletin*

- no. 16 International Crops Research Institute for the Semi-Arid Tropics (Patancheru), 36.
- Thakur, R. P., and Rai, K. N. (2002). "Pearl millet ergot research: advances and implications," in *Sorghum and Millets Diseases*, ed J. F. Leslie (Ames, IA: Iowa State Press), 57–64. doi: 10.1002/9780470384923.ch9
- Thakur, R. P., Sharma, R., and Rao, V. P. (2011). *Screening Techniques for Pearl Millet Diseases*. Information Bulletin No. 89. Patancheru: International Crops Research Institute for the Semi-Arid Tropics, 56.
- Thakur, R. P., and Williams, R. J. (1980). Pollination effects on pearl millet ergot. *Phytopathology* 70, 80–84. doi: 10.1094/Phyto-70-80
- Thakur, R. R., Rai, K. N., King, S. B., and Rao, V. P. (1993). "Identification and utilization of ergot resistance in pearl millet," in *Research Bulletin no. 17. International Crops Research Institute for the Semi-Arid Tropics* (Patancheru), 40.
- Tuberosa, R. (2012). Phenotyping for drought tolerance of crops in the genomics era. *Front. Physiol.* 3:347. doi: 10.3389/fphys.2012.00347
- Upadhyaya, H. D., Yadav, D., Reddy, K. N., Gowda, C. L. L., and Singh, S. (2011). Development of pearl millet mini core collection for enhanced utilization of germplasm. *Crop Sci.* 51, 217–223. doi: 10.2135/cropsci2010.06.0336
- Vadez, V., Kholová, J., Hummel, G., Zhokhavets, U., Gupta, S. K., and Hash, C. T. (2015). LeasyScan: a novel concept combining 3D imaging and lysimetry for high-throughput phenotyping of traits controlling plant water budget. *J. Exp. Bot.* 66, 5581–5593. doi: 10.1093/jxb/erv251
- van Oosterom, E. J., Biding, F. R., and Weltzien, R. E. (2003). A yield architecture framework to explain adaptation of pearl millet to environmental stress. *Field Crops Res.* 80, 33–56. doi: 10.1016/S0378-4290(02)00153-3
- van Oosterom, E. J., Mahalakshmi, V., Biding, F. R., and Rao, K. P. (1995). Effect of water availability and temperature on the genotype by environment interaction of pearl millet in semi-arid tropical environments. *Euphytica* 89, 175–183. doi: 10.1007/BF00034603
- Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nature Biotech.* 35, 969–974. doi: 10.1038/nbt.3943
- Varshney, R. K., Sinha, P., Singh, V. K., Kumar, A., Zhang, Q., and Bennetzen, J. L. (2020). 5Gs for crop genetic improvement. *Curr. Opin. Plant Biol.* 56, 190–196. doi: 10.1016/j.pbi.2019.12.004
- Velu, G., Rai, K. N., Sahrawat, K. L., and Sumalini, K. (2008). Variability for grain iron and zinc contents in pearl millet hybrids. *J. SAT Agric. Res.* 6, 1–4. Available online at: http://oar.icrisat.org/5043/1/SATeJou6_1-4_2008.pdf
- Witcombe, J. R. (1999). "Population improvement," in Pearl Millet Breeding, eds I. S. Khairwal, K. N. Rai, D. J. Andrews, and G. Harinarayana (New Delhi: Oxford and IBH), 213–256.
- Yadav, O. P. (1996). Downy mildew incidence of pearl millet hybrids with different male-sterility-inducing cytoplasm. *Theor. Appl. Genet.* 92, 278–280. doi: 10.1007/BF00223386
- Yadav, O. P. (2011). Breeding dual-purpose pearl millet (*Pennisetum glaucum*) for northwestern India: understanding association of biomass and phenotypic traits. *Indian J. Agric. Sci.* 81, 816–820.
- Yadav, O. P., Biding, F. R., and Singh, D. V. (2009). Utility of pearl millet landraces in breeding dual-purpose hybrids for arid zone environments of India. *Euphytica* 166, 239–247. doi: 10.1007/s10681-008-9834-y
- Yadav, O. P., Khairwal, I. S., and Singh, S. (1992). Smut severity of pearl millet hybrids with male sterile and fertile cytoplasm. *Euphytica* 64, 139–142.
- Yadav, O. P., Manga, V. K., and Gupta, G. K. (1993). Influence of A1 cytoplasmic substitution on the downy mildew incidence of pearl millet. *Theor. Appl. Genet.* 87, 558–560. doi: 10.1007/BF00221878
- Yadav, O. P., and Rai, K. N. (2011). Hybridization of Indian landraces and African elite composites of pearl millet results in biomass and stover yield improvement under arid zone conditions. *Crop Sci.* 51, 1980–1987. doi: 10.2135/cropsci2010.12.0731
- Yadav, O. P., and Rai, K. N. (2013). Genetic improvement of pearl millet in India. *Agric. Res.* 2, 275–292. doi: 10.1007/s40003-013-0089-z
- Yadav, O. P., Rai, K. N., Biding, F. R., Gupta, S. K., Rajpurohit, B. S., and Bhatnagar, S. K. (2012c). Pearl millet (*Pennisetum glaucum*) restorer lines for breeding dual-purpose hybrids adapted to arid environments. *Indian J. Agric. Sci.* 82, 922–927.
- Yadav, O. P., Rai, K. N., Khairwal, I. S., Rajpurohit, B. S., and Mahala, R. S. (2011). "Breeding pearl millet for arid zone of north-western India: constraints, opportunities and approaches," in All India Coordinated Pearl Millet Improvement Project (Jodhpur), 28.
- Yadav, O. P., Rai, K. N., Rajpurohit, B. S., Hash, C. T., Mahala, R. S., Gupta, S. K., et al. (2012b). "Twenty-five years of pearl millet improvement in India," in All India Coordinated Pearl Millet Improvement Project (Jodhpur), 122.
- Yadav, O. P., Rajpurohit, B. S., Kherwa, G. R., and Kumar, A. (2012a). Prospects of enhancing pearl millet (*Pennisetum glaucum*) productivity under drought environments of north-western India through hybrids. *Indian J. Genet. Plant Breed* 72, 25–30.
- Yadav, O. P., Singh, D. V., Dhillon, B. S., and Mohapatra, T. (2019). India's evergreen revolution in cereals. *Curr. Sci.* 116, 1805–1808. doi: 10.18520/cs/v116/i11/1805-1808
- Yadav, O. P., Upadhyaya, H. D., Reddy, K. N., Jukanti, A. K., Pandey, S., and Tyagi, R. K. (2017). Genetic resources of pearl millet: status and utilization. *Indian J. Plant Genet Resour.* 30, 31–47. doi: 10.5958/0976-1926.2017.00004.3
- Yadav, R. S., Hash, C. T., Biding, F. R., Cavan, G. P., and Howarth, C. L. (2002). Quantitative trait loci associated with traits determining grain and stover yield in pearl millet under terminal drought stress conditions. *Theor. Appl. Genet.* 104, 67–83.
- Yadav, R. S., Hash, C. T., Biding, F. R., Devos, K. M., and Howarth, C. J. (2004). Genomic regions associated with grain yield and aspects of post-flowering drought tolerance in pearl millet across stress environments and tester background. *Euphytica*, 136, 265–277.
- Yella Goud, T., Sharma, R., Gupta, S. K., Uma Devi, G., Gate, V. L., and Boratkar, M. (2016). Evaluation of designated hybrid seed parents of pearl millet for blast resistance. *Indian J. Plant Protect.* 44, 83–87.

Conflict of Interest: RM was employed by company SeedWorks International Pvt Ltd, Hyderabad, India.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Yadav, Gupta, Govindaraj, Sharma, Varshney, Srivastava, Rathore and Mahala. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Understanding the Sorghum–*Colletotrichum sublineola* Interactions for Enhanced Host Resistance

Kibrom B. Abreha*, Rodomiro Ortiz, Anders S. Carlsson and Mulatu Geleta

Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

OPEN ACCESS

Edited by:

Mahalingam Govindaraj,
International Crops Research Institute
for the Semi-Arid
Tropics (ICRISAT), India

Reviewed by:

Rasappa Viswanathan,
Indian Council of Agricultural
Research (ICAR), India
Rajan Sharma,
International Crops Research Institute
for the Semi-Arid
Tropics (ICRISAT), India

*Correspondence:

Kibrom B. Abreha
kibrom.abreha@slu.se

Specialty section:

This article was submitted to
Plant Pathogen Interactions,
a section of the journal
Frontiers in Plant Science

Received: 15 December 2020

Accepted: 22 March 2021

Published: 20 April 2021

Citation:

Abreha KB, Ortiz R, Carlsson AS and
Geleta M (2021) Understanding the
Sorghum–*Colletotrichum sublineola*
Interactions for Enhanced
Host Resistance.
Front. Plant Sci. 12:641969.
doi: 10.3389/fpls.2021.641969

Improving sorghum resistance is a sustainable method to reduce yield losses due to anthracnose, a devastating disease caused by *Colletotrichum sublineola*. Elucidating the molecular mechanisms of sorghum–*C. sublineola* interactions would help identify biomarkers for rapid and efficient identification of novel sources for host-plant resistance improvement, understanding the pathogen virulence, and facilitating resistance breeding. Despite concerted efforts to identify resistance sources, the knowledge about sorghum–anthracnose interactions remains scanty. Hence, in this review, we presented an overview of the current knowledge on the mechanisms of sorghum–*C. sublineola* molecular interactions, sources of resistance for sorghum breeding, quantitative trait loci (QTL), and major (*R*-) resistance gene sequences as well as defense-related genes associated with anthracnose resistance. We summarized current knowledge about *C. sublineola* populations and its virulence. Illustration of the sorghum–*C. sublineola* interaction model based on the current understanding is also provided. We highlighted the importance of genomic resources of both organisms for integrated omics research to unravel the key molecular components underpinning compatible and incompatible sorghum–anthracnose interactions. Furthermore, sorghum-breeding strategy employing rapid sorghum germplasm screening, systems biology, and molecular tools is presented.

Keywords: sorghum, anthracnose, quantitative trait loci, *R*-genes, germplasm, host-plant resistance, *Colletotrichum sublineola*

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] – a diploid photosynthesis efficient C_4 crop – is one of the most important cereals serving as a staple food for over 500 million people globally, used as animal feed, and is increasingly important source of biomass for cellulosic ethanol production. Due to its rich genetic diversity and adaptability to adverse conditions such as drought, nowadays sorghum is annually cultivated on over 42 million ha across six continents (FAOSTAT, 2020). In spite of its diverse use and resilience, sorghum is predominantly produced by subsistence farmers in the developing world (**Figure 1**) significantly contributing to food security. However, biotic and abiotic stresses are causing significant yield losses across all its growing areas.

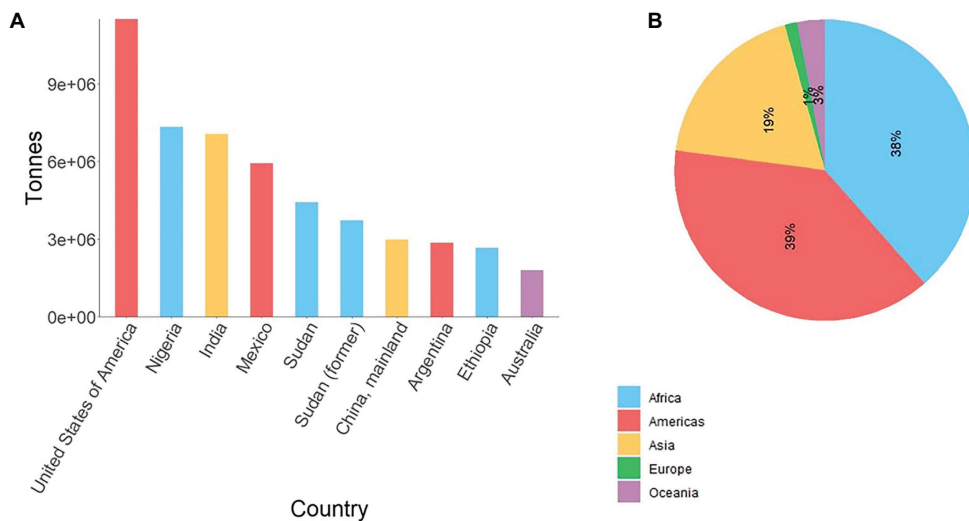


FIGURE 1 | (A) Average production of sorghum in top 10 sorghum producing countries from 1994 to 2018, and **(B)** corresponding production share by region. Data source (FAOSTAT, 2020). Sudan (former) reflects average sorghum production in Sudan and South Sudan up to 2011. For Sudan, the data represent average production from 2012 to 2018.

Diseases, such as stalk rot, downy mildew, grain mold, rust, head smut, leaf blight, and anthracnose are constraining global sorghum production (Wang et al., 2006; Little et al., 2012; Tesso et al., 2012; Das and Rajendrakumar, 2016; Mengistu et al., 2018). Because of its wide distribution and ability to infect all above ground parts of the plant, anthracnose caused by the destructive fungal pathogen *Colletotrichum sublineola* is one the most important diseases of sorghum. Since it was first reported in Togo in 1902 and later in United States in 1912 (Crouch and Tomaso-Peterson, 2012), the pathogen has spread to almost everywhere where sorghum is grown. It is causing significant yield losses annually, especially in the tropical and subtropical regions where there are favorable climatic conditions for disease development (Erpelding, 2008). Yield losses of up to 67% due to anthracnose are recorded in susceptible sorghum cultivars (Mengistu et al., 2018) but, without efficient disease management practices, the pathogen can devastate the whole crop.

Agronomic practices alone are not effective to reduce infections and yield losses due to anthracnose. Although fungicide application is an effective method to control the disease and reduce yield losses (Acharya et al., 2019), it is not economically and practically feasible for small-scale farmers and is not environmentally friendly. Growing sorghum cultivars resistant to anthracnose is considered the most efficient and is a core in integrated strategy for anthracnose management. To this end, improving anthracnose resistance has been an utmost priority in sorghum breeding programs. Fortunately, there is a high genetic diversity and wide anthracnose resistance variation in sorghum landraces (Motthaodi et al., 2017; Afolayan et al., 2019; Cuevas and Prom, 2020; Mengistu et al., 2020), which can be explored and used in its breeding program for improving resistance against the disease.

Traditional breeding is a relatively slow process and not sufficient to tap full potential of crop genetic resources.

The use of DNA markers may increase the pace and efficiency of plant breeding (Zheng et al., 2011; Upadhyaya et al., 2013; Kage et al., 2016; Ordonio et al., 2016). Accordingly, identifying naturally occurring major disease resistance genes (*R*-genes) in wild relatives and landraces of a crop and introducing these genes into elite materials using molecular tools are the most widely adopted strategy to develop disease-resistant cultivars of several crops. The sorghum genome contains hundreds of putative *R*-gene sequences (Mace et al., 2014; McCormick et al., 2018), some of which are localized within the anthracnose resistance associated quantitative trait loci (QTL) regions that were identified in several landraces (Upadhyaya et al., 2013; Felderhoff et al., 2016; Patil et al., 2017; Xu, 2019). These findings suggest the presence of rich sorghum genetic resource that can be used for enhancing sorghum anthracnose resistance. However, none of these *R*-genes have been functionally validated for resistance against *C. sublineola* and introduced into sorghum elite cultivars so far.

Deep understanding of the sorghum-*C. sublineola* interactions is needed in order to rapidly identify resistance sources in the diverse landraces, isolation, and characterization of the *R*-genes (Xu, 2019), and elucidate the virulence of the pathogen (Buiate et al., 2017). In line with this, improving our knowledge on molecular mechanisms of the interaction process can facilitate resistance cultivar development and design an effective disease management strategy against anthracnose. So far, little is known about the underlying molecular events underpinning the sorghum-*C. sublineola* interactions and the existing findings remain scattered. In this review, we present an overview of the current knowledge on molecular-level sorghum-*C. sublineola* interactions. This article summarizes the *R*-genes and QTL associated with the resistance as well other anthracnose defense-related genes in sorghum. It also provides a summary of current knowledge related to anthracnose population structure and

virulence, which would be crucial for the improvement of the crop's resistance against the pathogen. Further, it highlights the potential use of current advances in omics techniques to enhance the understanding of the interaction process and suggests future directions to accelerate the breeding of resistant sorghum cultivars.

SORGHUM GENETIC RESOURCES FOR ENHANCING ANTHRACNOSE RESISTANCE

Crop germplasm contains genetic makeup of a specific species. Phenotypic and genotypic diversity in a crop germplasm is a foundation for improving its agronomic traits, such as biotic and abiotic stress tolerance, yield, and nutritional quality. A large sorghum germplasm collection that has a paramount importance in sorghum breeding is available (Upadhyaya et al., 2016). Globally, over 236,000 sorghum accessions are maintained in several gene banks (Upadhyaya et al., 2016), of which more than 44,000 sorghum accessions originating from 114 countries, are conserved in the National Plant Germplasm System (NPGS) in United States (Cuevas et al., 2017). Although there could be some overlaps with the NPGS germplasm collection, the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India also maintains 39,923 sorghum accessions (Upadhyaya et al., 2017). Sorghum germplasm collections also exist in several countries that are recognized as its center of origin and/or diversity, such as Ethiopia, Eritrea, and Sudan (Girma et al., 2019; Cuevas and Prom, 2020; Mengistu et al., 2020). However, sorghum wild relatives that have a potential to serve as promising sources of genes for sorghum improvement (Ananda et al., 2020) remain neglected (Upadhyaya et al., 2016). The remarkable phenotypic and genetic diversity within these collections suggest presence of rich sorghum genetic resources. Phenotyping and genotypic characterization of the crop landraces and their wild relatives is a fundamental step for the identification of genotypes that serve as a novel source of resistance in sorghum breeding programs (Upadhyaya et al., 2016).

A diverse sorghum germplasm is crucial genetic resource for breeding programs aimed to improve sorghum anthracnose resistance (Cuevas et al., 2017). Aiming to identify novel sources of anthracnose resistance within the diverse sorghum germplasm, several large-scale mass resistance-screening assays have identified anthracnose resistant genotypes with potential use in sorghum breeding. Thakur et al. (2007) evaluated anthracnose resistance of 15 sorghum lines originating from six countries in 14 anthracnose hotspots in Africa and Asia for 4–7 years, and identified line IS 12467 and IS 6928 that potentially harboring different resistance genes and hence could be integrated into sorghum breeding programs. Anthracnose resistance evaluation of 87 sorghum lines and 63 hybrids against 12 isolates under field conditions shows vertical and horizontal resistance of the host (Buiate et al., 2010). These results indicate the presence of major and minor additive genes in the sorghum germplasm conferring race specific and

race nonspecific resistance, respectively. Developing and deploying cultivars with both types of resistance against anthracnose would enhance its efficacy and durability. Anthracnose resistant and susceptible accessions were identified in sorghum germplasm randomly selected from USDA-ARS NPGS (Prom et al., 2012a; Cuevas et al., 2014), Burkina Faso and South Africa (Cuevas et al., 2016), Ethiopia (Mengistu et al., 2019), and China (Xu et al., 2020). Resistance evaluation of sorghum germplasm originating from various geographical regions using several strains of anthracnose, collected from diverse agro-ecologies, across growing seasons would lead to the identification of a highly useful resistance sources for developing sorghum cultivars with wide adaptation. Such studies would diversify potential sources of resistance for sorghum breeding and facilitate efficient utilization of the germplasm. However, despite the efforts by Thakur et al. (2007), there is lack of concerted effort to evaluate accessions across contrasting environments, harboring different *C. sublineola* populations, to identify resistance sources useful for global sorghum breeding. Overall, resistance evaluation studies have a great potential to identify resistance sources that can be directly used in resistant cultivar development through conventional plant breeding or the application of molecular breeding tools, and are highly useful for understanding the inheritance of the anthracnose resistance in sorghum.

INHERITANCE OF ANTHRACNOSE RESISTANCE IN SORGHUM

Understanding the inheritance of anthracnose resistance in sorghum is a key to identify novel sources, quickly transfer the resistance into elite materials, and enhance durability of introduced resistance to cope up with the diverse and evolving pathogen populations. In line with this, a single dominant gene for resistance against *C. sublineola* was first identified in 1950 (Lebeau et al., 1950). The presence of QTLs associated with the resistance as well as wide variability of the resistance phenotype found in sorghum landraces suggest that anthracnose resistance is a multigenic trait (Cuevas et al., 2014; Ahn et al., 2019; Mengistu et al., 2019; Cuevas and Prom, 2020; Xu et al., 2020). Nevertheless, the anthracnose resistance in sorghum mostly exhibits dominant genetic inheritance even if it also segregates as a recessive trait (Mehta et al., 2005; da Costa et al., 2011).

Using F₃ population derived from a susceptible (BTx623) and resistant (SC326-6) parents, Boora et al. (1998) reported that the anthracnose resistance is inherited as a single gene recessive trait. Two random amplified polymorphic DNA (RAPD) markers, OPD 16 and OPD 12, were found linked to a recessive resistance allele in sorghum accessions (Panday et al., 2002). A sequence characterized RAPD marker OPJ 011437 is linked to the anthracnose resistance segregating as a recessive in sorghum (Singh et al., 2006). Identification and characterization of these recessive genes conferring resistance would be crucial in order to apply loss-of-function mutation strategies in their dominant homologs to achieve resistance against anthracnose. Although DNA markers found linked to

the recessive anthracnose resistance trait in sorghum genotypes are limited, several studies have identified markers associated with the resistance segregating as a dominant trait (Mohan et al., 2010; Burrell et al., 2015; Felderhoff et al., 2016; Cuevas et al., 2017). These markers linked to both forms of resistance trait are crucial resources for marker-assisted selection (MAS) of accessions with novel resistance and their utilization in molecular breeding for developing anthracnose resistant sorghum cultivars. Nevertheless, despite the presence of draft genome sequence of sorghum for over a decade (Paterson et al., 2009), the identified markers have not been validated and utilized in sorghum molecular breeding. This is due to limited efforts to improve the genome annotation, with only few genotypes whole genome-sequenced so far to capture genome-wide sequence polymorphism. Moreover, most of the identified markers are linked to QTLs making it challenging to validate them across environments.

It is worth noting that some sorghum genotypes displaying dominant resistance could also harbor additional resistance genes that show recessive inheritance. Hence, molecular characterization of the variable resistance phenotypes in sorghum would help identify genotypes containing only a dominant resistance gene, or a recessive gene, or both. Classification of the diverse landraces into the three groups, based on levels of disease resistance, is helpful for efficient utilization of the gene pool for sorghum resistance improvement. For instance, since deploying the dominant resistance may increase virulence alleles of the pathogen, augmenting resistance with recessive resistant genes in breeding programs may enhance durability of the anthracnose resistance. However, most sorghum breeding efforts so far have focused on the identification of dominant resistance, mostly conditioned by major resistance (*R*-) genes, because its effect is easier to characterize phenotypically.

MOLECULAR ASPECTS OF SORGHUM-*C. sublineola* INTERACTIONS

Plants are continuously interacting with myriad of microbes and have evolved mechanisms to fend off pathogen attack. The plant response to pathogen infection is generally described by the co-evolutionary zigzag model of plant-microbe interactions (Jones and Dangl, 2006). Describing the complex sorghum-*C. sublineola* interactions within the frame of the zigzag model would help illustrate the current understanding of the pathosystem and elucidate molecular basis of the host resistance and virulence of the pathogen during different phases of the interaction processes.

The *C. sublineola* infection processes on sorghum leaves and disease development have been previously described using cytological and ultrastructural studies (Wharton and Julian, 1996; Wharton et al., 2001), and elegantly reviewed by Crouch and Beirn (2009). *C. sublineola* is a hemibiotrophic pathogen that requires an initial biotrophic phase and, if the infection is successful, transition to a necrotrophic phase happens at ca. 66 h after inoculation (Wharton et al., 2001).

Although it remains to be elucidated, the transition could be facilitated by *C. sublineola* plant cell wall (CW) degrading enzymes (CWDEs; Wharton and Julian, 1996; Wharton et al., 2001) to facilitate the infection process. *Sorghum* responds to this initial stages of the *C. sublineola* infection by cell wall apposition forming papillae and accumulation of polyphenolics, phytoalexins, callose, hydrogen peroxide (H₂O₂), and hydroxyprolinerich glycoproteins (HRGPs; Basavaraju et al., 2009). Several studies also reported higher expression of several defense related genes encoding pathogenesis-related protein 10 (PR10), chitinase (PR3), and chalcone synthase, and thaumatin-like-protein (TLP; Lo et al., 1999; Li et al., 2013; Ahn et al., 2019). These observed defense responses against pathogen infection are hallmarks of the host basal defense response termed as pathogen/microbe-associated molecular patterns (PAMPs/MAMPs) triggered immunity (PTI) reported in several plants (Jones and Dangl, 2006). PTI is activated when the PAMPs, conserved molecular signatures in the cell wall, are recognized by the pattern recognition receptors (PRRs) on the host cell surface (Jones and Dangl, 2006). The presence of enhanced chitinase (PR3) and β -1,3-glucanase in sorghum plants inoculated with *C. sublineola* (Li et al., 2013; Ahn et al., 2019), for instance, may indicate that chitin and β -glucan in the pathogens' cell wall serve as PAMPs (Fesel and Zuccaro, 2016) and have been detected by the host receptors and triggered defense response. Biochemical studies characterizing *C. sublineola* cell wall composition and sorghum receptors, which recognize the pathogen and activate host defense response, would provide crucial insights into early molecular events of the interaction process.

The basal defense response against anthracnose appears similar in the resistant and susceptible sorghum genotypes (Wharton et al., 2001). However, the pathogen can successfully penetrate the papillae, invade adjacent host cells, continue colonizing host tissue, and develops a disease in the susceptible genotypes (Wharton and Julian, 1996; Wharton et al., 2001; Basavaraju et al., 2009). The presence of small-secreted proteins, some of which annotated as effectors related to virulence of the pathogen, in the *C. sublineola* genome (Buiate et al., 2017) suggests that the pathogen could secrete plethora of these proteins to manipulate the host defense. Such proteins evading the host response are known as virulence (*avr*) effectors triggering susceptibility (ETS; Jones and Dangl, 2006). In contrast, hypersensitive response around the infection site due to accumulation of hydrogen peroxide and localized programmed cell death reinforces the PTI defense in the resistant sorghum genotypes (Wharton and Julian, 1996; Basavaraju et al., 2009). The defense response in resistant sorghum genotypes significantly reduces the germination of conidia, formation of appressorium, blocks penetrations, disrupts the pathogen cells, and restricts the pathogen growth at the biotrophic phase (Basavaraju et al., 2009; Tugizimana et al., 2018, 2019). Stronger basal defense response in resistant genotypes could be attributed to earlier accumulation, and higher quantity and diversity of the phytoalexins and expression of defense-related genes (Lo et al., 1999; Basavaraju et al., 2009), probably conditioned by higher efficiency of pathogen detection and signal transduction in

the resistant genotypes. In addition, stronger reaction in the resistant genotypes could be related to recognition of the *C. sublineola* secreted proteins, known as avirulence (Avr) effectors, by the host nucleotide binding (NB)-leucine rich repeat (LRR; NB-LRR) encoding genes (Jones and Dangl, 2006). The recognition activates the effector triggered immunity (ETI), characterized by elevated accumulation of H_2O_2 leading to a programmed cell death around the point of infection that suppresses pathogen growth (Jones and Dangl, 2006).

The host NB and LRR domain containing proteins in the resistant genotypes recognize corresponding Avr effectors, described in gene-for-gene interaction. However, the defense response in sorghum involves several other genes as exemplified by a high degree of anthracnose resistance variability among the landraces and the QTLs related to *C. sublineola* resistance (Lebeau et al., 1950; da Costa et al., 2011; Patil et al., 2017; Prom et al., 2018; Ahn et al., 2019). An overview of the

interactions involving *C. sublineola* infection and sorghum defense response is presented in Figure 2.

MAJOR RESISTANCE GENES CONFERRING RESISTANCE AGAINST *C. sublineola*

Major resistance (*R*-) genes have evolved in plants that directly or indirectly recognize corresponding pathogen secreting small molecules known as effectors. The recognition activates the host resistance response termed ETI. Most of the *R*-genes identified so far in several plants encode proteins containing a central NB and C-terminal LRR domains. Based on their N-terminal domain, NB-LRR encoding genes are grouped into different classes (Mace et al., 2014), but most of the functionally

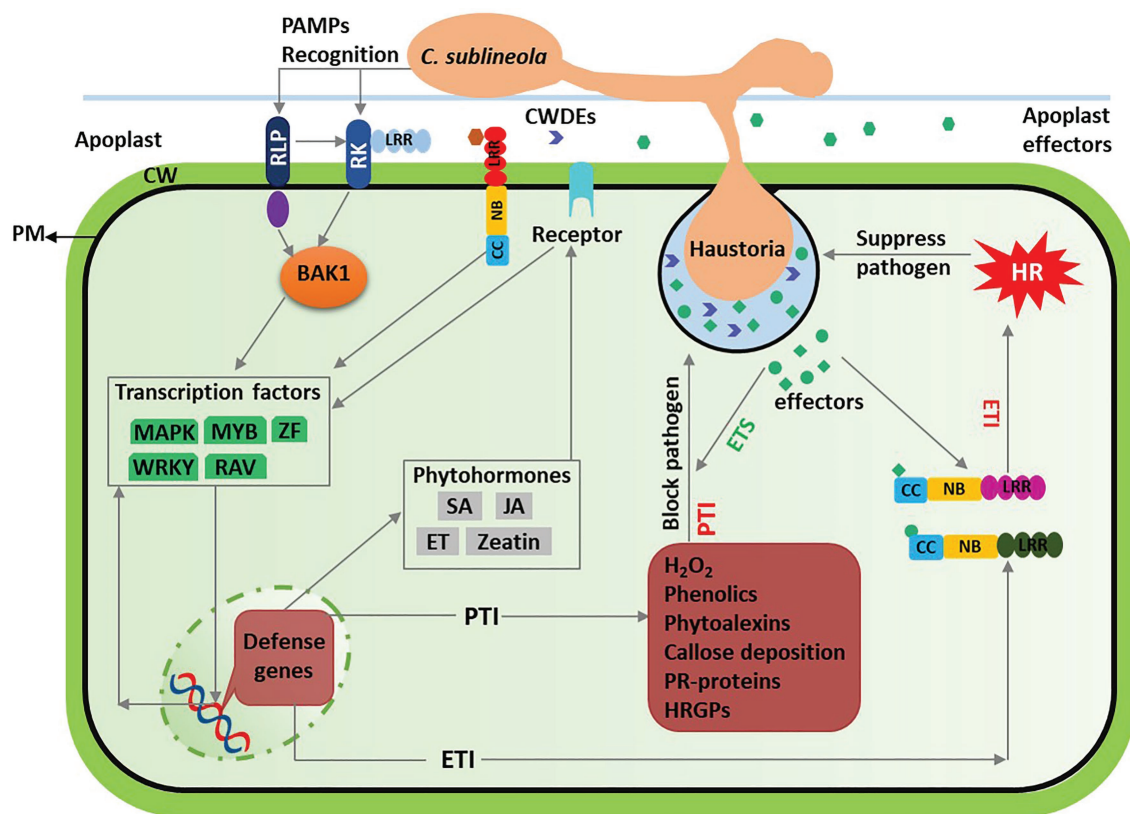


FIGURE 2 | Overview of sorghum-*Colletotrichum sublineola* interactions. Spores of *C. sublineola* land on sorghum tissue, propagate into germ tubes, and form appressoria, which are specialized infection structures. The infection penetrates the host cell wall (CW) and forms feeding structures called haustoria. During the course of infection, pathogenicity molecules, such as cell wall degrading enzymes (CWDEs) and effectors are secreted into the extracellular space. The CWDEs and ubiquitous structural molecules such as chitin are recognized as Pathogen associated molecular patterns (PAMPs) by the host cell receptors, receptor kinases (RK), and receptor like proteins (RLP). These receptors interact with extracellular leucine rich repeat (LRR) and intracellular BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE 1 (BAK1) activating the PAMP-triggered immunity (PTI). Host PTI response is characterized by increased accumulation of hydrogen peroxide (H_2O_2), phenolics, phytoalexins, hydroxyprolinerich glycoproteins (HRGPs), pathogenesis-related (PR-) proteins, and callose deposition around the point of infection and suppress pathogen growth. In susceptible genotypes, the pathogen secretes effectors that suppress the PTI whereas in resistant genotypes these extracellular and intracellular effectors are recognized by the host nucleotide-binding (NB) LRR (NB-LRR) receptors. This recognition induces the effector-triggered immunity (ETI) primarily recognized by accumulation of H_2O_2 resulting in hypersensitive response (HR), which is a form of programmed cell death. The ETI response blocks *C. sublineola* transition to necrotrophic phase and arrests the pathogen growth. Both PTI and ETI responses involve transcriptional factors (green) and phytohormones (gray), involved in signaling transduction, regulation of defense-related genes, and host physiology.

validated *R*-genes in plants so far belong to the coiled-coil domain containing NB-LRR (CC-NB-LRR) type.

Using the sorghum genome sequence (Paterson et al., 2009), Mace et al. (2014) identified more than 346 NB-encoding genes across the 10 chromosomes of the ca 700 Mbp genome. The most recent reannotation that has incorporated additional ca 30 Mbp of sequence has increased the number of genes in the sorghum genome by 24% to 34,211 (McCormick et al., 2018). If these additional sequences are included in the analysis, the number of NB-encoding genes could be even higher. Moreover, recent advancements in plant *R*-gene profiling using single-molecule real-time sequencing of resistance genes (SMRT-Seq; Jupe et al., 2014; Witek et al., 2016; Parra et al., 2019) and improving annotation tools like the nucleotide-binding and leucine-rich-repeat NLR-annotator (Steuernagel et al., 2020) may lead to the identification of more *R*-genes in sorghum. Relative to other gene families, the NB-LRR gene family in sorghum shows higher diversity (Mace et al., 2014). This may suggest that a key component of a defense response against fast-evolving pathogens, the NB-LRR genes, is under selection pressure. Interestingly, about 80% of the nucleotide binding site (NBS)-encoding genes are located within the QTL regions of the sorghum genome controlling fungal pathogen resistance (Mace et al., 2014) suggesting that several of these genes could be related to anthracnose resistance in sorghum.

Several studies identified anthracnose resistance QTLs harboring NB-LRR genes in sorghum (Upadhyaya et al., 2013; Felderhoff et al., 2016; Patil et al., 2017; Cuevas et al., 2018, 2019). A sorghum LRR encoding gene associated with anthracnose resistance was found in a locus located on chromosome 6 (Mohan et al., 2010). Using 14,739 SNP markers, Upadhyaya et al. (2013) mapped eight loci linked to anthracnose resistance in sorghum and identified two NB-coding genes on chromosome 10. Genes encoding NB-LRR proteins were among disease resistance genes found in an anthracnose resistance-related QTL located on chromosome 5 (Burrell et al., 2015). Felderhoff et al. (2016) identified NB-LRR protein encoding genes among defense-related genes found within anthracnose resistance loci on chromosome 7 and 9. Patil et al. (2017) detected four NB-LRR genes within a QTL region on chromosome 9 conferring resistance against anthracnose. Cuevas et al. (2018, 2019), respectively identified loci containing CC-NBS-LRR genes in the United States sorghum association panel and Ethiopian sorghum accessions at NPGS. Two NBS-LRR encoding genes, Sobic.008G166400 and Sobic.008G166550, were identified in the QTL region at distal end of chromosome 8 segregating with *C. sublineola* resistance in a sorghum crossing population (Xu, 2019). The known NB-LRR sequences associated with anthracnose resistance in sorghum are summarized in **Table 1**. Identifying markers closely linked to the QTL regions and developing specific primers for these NB-LRR gene sequences are an important step for efficient identification of landrace genotypes for sorghum breeding. Resistant genotypes with other desirable agronomic traits such as seed yield can be directly deployed for production, or used as a source of resistance trait for improving elite materials using marker-assisted conventional

breeding as well as transgenic methods. The resistance trait in the sources genotype can be boosted using genome-editing methods. Therefore, functional validation of the *R*-genes would lead to efficient utilization of the genetic resources for anthracnose resistance improvement.

A study by Biruma et al. (2012) showed that silencing of two NB-LRR encoding genes, the *Cs1A* and *Cs2A* identified using cDNA-amplified fragment length polymorphism (AFLP) transcript profiling of resistant and susceptible sorghum genotypes, promote the anthracnose infection (Biruma et al., 2012). Moreover, the *SbLRR2*, a gene encoding a simple extracellular LRR protein, showed stronger expression in anthracnose resistant sorghum genotypes compared to susceptible genotypes, suggesting its role in the resistance against anthracnose (Zhu et al., 2015). The NB-LRR genes from common bean (Wu et al., 2017), tea (Shi et al., 2016), strawberry (Liab et al., 2013), and *Arabidopsis* (Birker et al., 2009) were shown to confer resistance against anthracnose in these crops. While these studies showed the role of NB-LRR genes in sorghum anthracnose resistance, their functional validation should confirm if these genes would maintain their function when transferred into a susceptible genotype. Several other sorghum NB-LRR genes need to be isolated, cloned, and transferred into susceptible genotypes and tested if they confer resistance against diverse strains of the pathogen. However, NB-LRR works in gene-for-gene interaction, recognizing cognate effector protein secreted by a strain of the pathogen (van der Biezen and Jones, 1998; Jones and Dangl, 2006). This indicates several NB-LRR and other defense-related genes are required for efficient plant resistance. Hence, isolation and functional characterization of several NBS-LRR gene sequences, localized with the QTLs conferring anthracnose resistance, would increase the *R*-gene resource availability for enhanced anthracnose resistance.

Several sorghum NB-LRR genes were also identified conferring resistance against multiple pathogens. For instance, the sorghum *SbLRR2* that is involved in anthracnose resistance conferred resistance against necrotrophic pathogens *Botrytis cinerea* and *Alternaria brassicicola* in transgenic *Arabidopsis thaliana* (Zhu et al., 2015). This may imply that the sorghum NB-LRR genes that confer resistance to fungal pathogens such as the *Setosphaeria turcica* (Martin et al., 2011) and *Periconia circinata* (Nagy and Bennetzen, 2008) could also confer resistance against anthracnose in sorghum. NB-LRR encoding genes cloned from sorghum and introduced to susceptible rice lines provide resistance to blast disease caused by *Magnaporthe oryzae* (Yang et al., 2013). The sorghum genome also contains homologous genes conferring resistance to common leaf rust (*Rp1*; NB-LRR) in maize (Chavan et al., 2015) and stripe rust resistance (*Yr10*; CC-NB-LRR) in wheat (Liu et al., 2014). Cloning of these genes from anthracnose resistant sorghum genotypes, introducing them to susceptible genotypes followed by their evaluation for anthracnose resistance is vitally important in order to consider their use in sorghum breeding. Such investigation could identify if any of the NB-LRR genes are involved in resistance against multiple pathogens posing threat in sorghum production. However, the plant host resistance does not rely only on the NB-LRR genes to thwart pathogen attack. Hence, the identification of other gene families

TABLE 1 | The list of quantitative trait loci (QTL) and NB-LRR genes identified in sorghum genotypes responding to anthracnose infection.

QTL/Gene	R-genes (NB-LRR)	Gene identification	Chromosome	Chromosomal position (Mbp)
QTL/Cs1A ¹	NB-LRR	Sb09g027470	SBI-09	4.9–5.05
QTL/Cs2A ¹	CC-NB-LRR	Sb09g004240	SBI-09	56.5–56.6
QTL/Cs1B ¹	CC-NB-LRR	Sb09g027520	SBI-09	4.9–5.1
QTL/Cs2B ¹	CC-NB-LRR	Sb09g004210	SBI-09	56.5–56.6
QTL ¹	CC-NB	Sb09g004215	SBI-09	4.9–5.05
QTL ¹	NB-LRR	Sb09g004220	SBI-09	4.9–5.05
QTL ¹	CC-LRR	Sb09g004230	SBI-09	4.9–5.05
QTL ²	LRR	Sobic.005G182400	SBI-05	60–72
QTL ²			SBI-01	60–72
QTL ²	CC-NB-LRR	Sobic.009G013300	SBI-09	–1.9
	LRR	Sobic.009G012900	SBI-09	–1.9
QTL ³	NB-LRR	Sobic.007G085400	SBI-07	0–55
QTL ³			SBI-09	0.5–3.5
QTL ⁴			SBI-04	0–11.4
QTL ⁴			SBI-06	0–6.3
QTL ⁴			SBI-06	39–40.9
QTL ⁴			SBI-06	45.2–49
QTL ⁵	CC-NB-LRR	Sobic.005G167500	SBI-05	63.68–65.66
	CC-NB-LRR	Sobic.005G167600	SBI-05	63.68–65.66
	CC-NB-LRR	Sobic.005G183000	SBI-05	63.68–65.66
	CC-NB-LRR	Sobic.005G183300	SBI-05	63.68–65.66
QTL ⁵	NB-ARC	Sobic.009G013000	SBI-09	57.42–58.36
	NB-ARC	Sobic.009G013100	SBI-09	57.42–58.36
	NB-ARC	Sobic.009G013300	SBI-09	57.42–58.36
Cg1 locus ⁶			SBI-05	
QTL ⁷	NB-ARC	Sb10g021850	SBI-10	48–48.65
	NB-ARC	Sb10g021860	SBI-10	48–48.65
SbLRR2 ⁸	LRR	Sb05g018800	SBI-05	55.03–55.04
QTL ⁹	NBS-LRR	Sb05g026470	SBI-05	53.80–62.15
	NBS-LRR	Sb05g026480	SBI-05	53.80–62.15

NB, nucleotide binding.

¹Biruma et al. (2012).

²Cuevas et al. (2018).

³Felderhoff et al. (2016).

⁴Mohan et al. (2010).

⁵Patil et al. (2017).

⁶Perumal et al. (2009).

⁷Upadhyaya et al. (2013).

⁸Zhu et al. (2015).

⁹Burrell et al. (2015).

involved in the plant defense response against the pathogen is equally relevant to facilitate breeding efforts.

SIGNALING CASCADES AND DEFENSE RESPONSE AGAINST ANTHRACNOSE

Plant defense response involves recognition of the pathogen, signaling transduction, and the response resulting from differential expression of several genes, proteins, and metabolites. The plant defense response commences when host receptors recognize the pathogen PAMPs and effectors. Due to the complexity of the defense response, several defense-related genes are involved in defense against pathogens. Although it does not fully demonstrate their function in anthracnose resistance, several genes, such as F-box domain, peroxidases, and Glucuronosyl transferases, chitinases, germinlike proteins, polyphenol oxidases, peroxidases, ABC-transporters, defensins, and related hypersensitive response have been identified within

the loci containing NB-LRR genes (Upadhyaya et al., 2013; Felderhoff et al., 2016; Cuevas et al., 2018). Using suppression subtractive hybridization (SSH), Li et al. (2013) identified several genes involved in signal transduction, secondary metabolism, protein synthesis, and degradation activated in response to *C. sublineola* infection of sorghum genotypes.

Sorghum Receptors Recognizing *C. sublineola*

Pattern-recognition receptors, such as receptor kinases (RKs) or receptor-like proteins (RLPs), are surface-localized receptors perceiving the pathogen PAMPs and trigger host defense response PTI (Zipfel, 2014). The ETI relies on R-receptors, encoded by the NB-LRR genes, which recognizes the pathogen secreted virulence effectors directly or indirectly (Zipfel, 2014). Interestingly, two genes annotated as cysteine-rich RLKs were identified in the NB-LRR genes containing anthracnose resistance associated QTL region on chromosome 5 in sorghum

(Patil et al., 2017). RLPs interact with LRR, and LRR-RLP complex requires the BRI1-ASSOCIATED KINASE-1 (*BAK1*) also known as *SERK1* involved in downstream signaling (Liebrand et al., 2014). Due to its broad role in several pathosystems, *BAK1* is considered central regulator of plant innate immunity (Zipfel, 2014). Yazawa et al. (2013) showed that *BAK1* is involved in resistance against a necrotrophic fungal pathogen *Bipolaris sorghicola* causing target leaf spot in sorghum. Although not experimentally validated so far, these results indicate a potential role of the RLKs-LRRs-*BAK1* complex in sorghum resistance against anthracnose. Following the pathogen recognition and activation of plant defense, transcription factors regulate expression of the genes involved in downstream signaling and response against the pathogen.

Transcription Factors Regulating Sorghum Response Against Anthracnose

The host plant defense signaling involves transcription factors that control the expression level of target genes and regulate cellular processes conditioning plant response to stresses (Pandey and Somssich, 2009). Three well-known family of transcription factors, the MYB, WRKY, and MAPK are involved in plant defense against various pathogens (Pandey and Somssich, 2009; Ambawat et al., 2013; Meng and Zhang, 2013). The sorghum genome contains 94 putative WRKY transcription factors (Baillio et al., 2020) as well as 128 MYB and 83 MAPK gene ontologies, found using keyword search in Phytozome v12.1 (Accessed on December 14, 2020). Some transcription factors are involved in sorghum resistance to anthracnose. A yellow seed1 (*y1*) encoding MYB transcription factor (Chopra et al., 2002; Boddu et al., 2005; Ibraheem et al., 2010) and a putative MYB domain protein 48 (*MYB48*; Patil et al., 2017) regulate the accumulation of phytoalexins and flavonoids in response to anthracnose infection in sorghum. Both antimicrobial compounds accumulate at the site of anthracnose infection (Lo et al., 1999; Basavaraju et al., 2009; Dao et al., 2011). In their recent study, Tugizimana et al. (2018) showed that phenylpropanoid and flavonoid pathways are central hub of the metabolism producing anti-fungal compounds in response to anthracnose infection in sorghum. Li et al. (2013) identified two MAPK transcription factors related genes in anthracnose infected sorghum seedlings. Downregulation of a zinc finger-like transcription factor (*Sb03g041170*) compromised *C. sublineola* resistance in sorghum (Biruma et al., 2012). A RAV transcription factor (*Sb01g049150* in locus 3) involved in *R*-gene resistance mediated pathway (Upadhyaya et al., 2013). Hence, functionally diverse transcriptional factors are involved in the sorghum defense response against anthracnose infection, which is characterized by changes in phytohormone levels (Tugizimana et al., 2018).

Role of Phytohormones in Anthracnose Resistance

Phytohormones, such as salicylic acid (SA), jasmonic acid (JA), and ethylene (ET), play crucial role in signaling pathways related to plant defense response against biotic and abiotic stresses (Verma et al., 2016). Accumulation of phytohormones, in response to pathogen infection, is perceived by receptor

proteins and subsequently initiates intracellular signal transduction and interacts with the transcription factors (Li et al., 2020). A recent study by Tugizimana et al. (2018) revealed quantitative changes of JA, SA conjugates, and abscisic acid (ABA) in anthracnose-infected sorghum. Sorghum genotypes with enhanced levels of amino acids (tyrosine, tryptophan), JA and SA conjugates, and zeatin were more resistant to *C. sublineola* (Tugizimana et al., 2018). Higher expression of chalcone synthase in *C. sublineola* inoculated sorghum (Ahn et al., 2019) could be associated with the accumulation of flavonoid and isoflavonoid phytoalexins and is involved in the salicylic acid mediated defense pathway, as reviewed in Dao et al. (2011). Likewise, JA (but not SA and ET) induced expression of the resistance gene *SbLRR2*, suggesting that it is involved in JA-mediated sorghum defense against the pathogen (Zhu et al., 2015). It is possible that the other resistance genes could be involved in SA and ET mediated defense response against the pathogen. Induction of ABA-responsive genes only in anthracnose inoculated resistant sorghum cultivars (Li et al., 2013) indicates role of the phytohormone in the resistance against the pathogen. However, the interactions among these phytohormones (resulting in synergy or antagonism) have not been discussed in relation to anthracnose resistance in sorghum. Plant response involving the recognition, signaling transduction, transcriptional regulation, phytohormone accumulation, and production of antimicrobial compounds is fine-tuned to the virulence of the infecting pathogen strain. This warrants knowledge about biology of the pathogen populations, pathogenicity, and virulence of its strains. This knowledge is crucial to understand the plant-defense responses and design optimized breeding programs for improving sorghum anthracnose resistance.

Colletotrichum sublineola CAUSING ANTHRACNOSE IN SORGHUM

Colletotrichum sublineola, causal agent of the devastating anthracnose disease in sorghum, is a hemibiotrophic fungal pathogen (Crouch and Beirn, 2009; Tesso et al., 2012). Previously, *C. sublineola* was known as *Colletotrichum graminicola*, but rDNA sequences, restriction fragment length polymorphism (RFLP) of mitochondrial DNA, mating analysis, and appressorial morphology revealed that these two species are distinctly different (Vaillancour and Hanau, 1992; Sherriff et al., 1995; Wharton and Julian, 1996). *C. sublineola* reproduces asexually but pairing of isolates induces teleomorph stage of sexual reproduction indicating that the pathogen is heterothallic (Vaillancour and Hanau, 1992). The presence of parasexual processes contributing to high genetic diversity in *C. sublineola* populations was also reported (Chala et al., 2011). This pathogen primarily overwinters as mycelium, acervuli, and sclerotia, alternatively, in the soil, seeds, and decomposing crop residues (Casela and Frederiksen, 1993; Crouch and Beirn, 2009). This indicates that the pathogen is well-adapted to seasonal changes and can easily transmit inoculum to new plants in the subsequent growing season. During favorable conditions, asexual conidia are

formed, disseminated through water splash, attached to the plant tissue, and cause a disease (Crouch and Beirn, 2009).

Understanding the life history traits of *C. sublineola* is fundamental for efficient identification of resistance sources and eventual introgression of resistance genes into elite cultivars. The knowledge about the pathogen's biology and lifecycle is also crucial for designing enhanced disease management strategies. An overview of the pathogen's biology and lifecycle including its morphology, overwintering and primary inoculum sources, plant-to-plant dissemination mechanisms, and reproductive structures, infection strategy was previously published (Crouch and Beirn, 2009). Elucidating genetic diversity and pathogenicity of *C. sublineola* populations, characteristics related to pathogen's survival and ability to cause a disease, and its evolutionary potential across major sorghum growing regions is paramount for developing durable resistance in sorghum, as presented in the following sections.

High-Diversity and Pathogenicity of *C. sublineola* Populations

Amplified fragment length polymorphism analysis revealed the presence of diverse *C. sublineola* isolates within a local field in Ethiopia (Chala et al., 2011; Chala, 2013). Likewise, 13 pathotypes were identified among 87 isolates of the pathogen collected from research stations and production fields in Arkansas, using differential reaction in sorghum genotypes (Moore et al., 2008). The results indicate a high diversity of the pathogen even within a single field. On the other hand, Rosewich et al. (1998) found not much variation within 411 *C. sublineola* isolates sampled from 1991 to 1993 from a sorghum disease nursery in Georgia, United States. As these authors noted, the presence of one predominant haplotype leads to low diversity of the pathogen population in the area. Variability of the pathogen is a significant challenge because it enables the pathogen to adapt to the deployed resistance quickly. Conducive environment, mutation (particularly in the effector proteins), and genetic recombination due to mating, epidemiology, and co-evolution with alternative hosts such as wild relatives as well as resistance deployment can increase pathogen aggressiveness or evolve new virulent strains and cause significant damage. Thus, the diversity of the pathogen within one field warrants breeding cultivars with a broad resistance that can cope with diverse and fast-evolving pathotypes. In addition, introduced resistance should be optimal for diverse sorghum growing areas with contrasting environmental conditions. Unfortunately, research showed that the anthracnose population structure is different across geographically distinct regions and varies across crop growing seasons. Using AFLP analysis, a high level of genetic variation among 102 isolates and genetic differentiation between isolates from different sites in Ethiopia were reported (Chala et al., 2011). More than 230 *C. sublineola* isolates collected from Texas, Arkansas, Georgia, and Puerto Rico in United States, characterized using AFLP markers exhibited a high genetic diversity and differential virulence profiles (Prom et al., 2012b). A RAPD based study revealed a high genetic variation among 37 *C. sublineola* isolates, which were clustered according to their geographic origin in Brazil (Valerio et al., 2005).

The virulence characterization of these isolates using differential sorghum genotypes led to the identification of 22 races (Valerio et al., 2005), indicating high diversity of the pathogen in genetic make-up and pathogenicity. Based on the resistance reactions of 15 sorghum inbred lines grown in 14 anthracnose hotspots in Asia and Africa, Thakur et al. (2007) found differences in virulence of the *C. sublineola* populations across the locations and growing seasons. Despite these efforts, given the wide distribution of the pathogen and its diversity, the pathogen populations across sorghum growing regions remains poorly understood. This needs globally coordinated efforts to develop robust genetic markers, set of sorghum differentials for testing pathogenicity as previously suggested (Tesso et al., 2012); and identify and characterize key molecular determinants of the pathogenicity such as the effector proteins (Buiate et al., 2017).

Deciphering the Role of Effector Proteins in *C. sublineola* Pathogenicity

Plant pathogens secrete effector proteins to overcome host defense response and manipulate cell physiology to cause a disease. The whole genome sequences of the pathogen followed by comparative genomics can provide crucial insights into the pathogen evolution and pathogenicity. Such approach also facilitates the development of molecular markers for the identification of different strains of the pathogen. The published *C. sublineola* genome (Baroncelli et al., 2014) contains genes coding for hundreds of secreted small molecules with 180 of them putatively annotated as effectors (Buiate et al., 2017) using the EffectorP prediction tool (Sperschneider et al., 2018). The number of predicted effectors in *C. sublineola* is higher compared to that of its closely related species *C. graminicola* (Buiate et al., 2017), thereby suggesting their more diverse role during the infection process. Similar to the work done in *C. graminicola* effectors in maize (Gong et al., 2020), bioinformatics analysis, isolation, and characterization of the *C. sublineola* effectors, would play a crucial role in elucidating the role of these effectors in determining the virulence of the pathogen and infection process in sorghum. Moreover, the cloning of *C. sublineola* effectors would facilitate future resistance screening of sorghum landraces and wild relatives using effectomics, a method for quick identification of major *R*-genes as shown in potato wild relatives (Domazakis et al., 2017). Characterizing the *C. sublineola* effectors would help efficiently determine the races of the pathogen and follow virulence pattern of *C. sublineola* populations across geographical regions.

TOWARDS OMICS-AIDED ANTHRACNOSE RESISTANCE IMPROVEMENT IN SORGHUM

The sorghum-anthracnose interaction is a complex process, which activates defense response characterized by the change in the expression of several genes (Li et al., 2013), proteins, and metabolites in the host plant (Tugizimana et al., 2018, 2019) and the pathogen. Advances in genomics, transcriptomics,

proteomics, and metabolomics studies can help elucidate the cellular processes during the host-pathogen interactions. Following the sequencing of the sorghum reference genome in 2009 (Paterson et al., 2009), several sorghum genotypes have been re-sequenced to differentiate and characterize their agronomic traits (Zheng et al., 2011; Evans et al., 2013) and understand the complex domestication history of the crop (Mace et al., 2013). Likewise, Baroncelli et al. (2014) reported the draft genome sequence of *C. sublineola*. These sorghum and *C. sublineola* genetic and genomic resources are important to understand their biology, identify and understand the genes underlying variation leading to resistance and pathogenicity, and hence facilitate genetic improvement of the crop (Buiate et al., 2017; Boyles et al., 2019). For example, the genome sequence data of both organisms were used for the identification of NB-LRR and other defense-related genes (Biruma et al., 2012; Mace et al., 2014) as well as understanding host-specificity and putative pathogenicity genes in *C. sublineola* (Buiate et al., 2017). Re-sequencing of several other genotypes and strains, and continuous improvement of genome annotation would facilitate

the use of functional genomics to understand the pathosystem and enhance sorghum resistance against the pathogen.

Existing genome sequence databases for both organisms make it possible to apply integrated omics approaches to understand their interactions, identify resistance sources, and facilitate the development of sorghum cultivars resistant to anthracnose through the application of advanced plant breeding methods (Figure 3). Not all genes contained in the genome but those transcriptionally induced, at particular time during compatible and incompatible sorghum-anthracnose interactions, lead to alterations in myriad of defense-response or pathogenicity related proteins and metabolites. Large-scale transcriptomics, proteomics, and metabolomics studies can capture these changes in the host and pathogen. Increasingly cheaper sequencing technologies and advancing instrumentation, data processing tools, and bioinformatics software make omics studies possible methods in deciphering sorghum-anthracnose interactions. However, it is only recently that transcriptomics (Wang et al., 2020) and metabolomics (Tugizimana et al., 2018) approaches are applied in sorghum-*C. sublineola* interactions. Since several

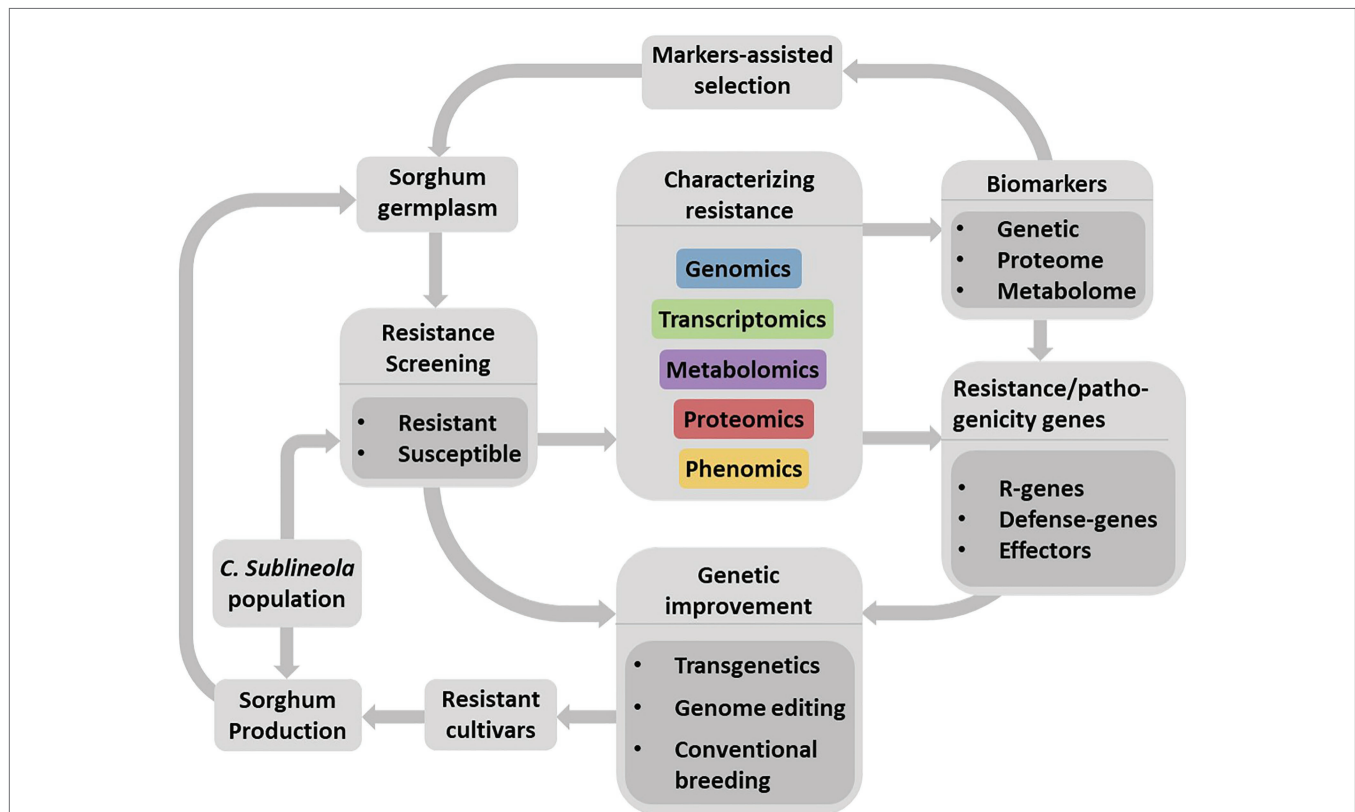


FIGURE 3 | Schematic flow representing a strategy for the development of anthracnose resistant sorghum cultivars through the identification of resistance sources, development and application of biomarkers as well as understanding the pathosystems. Large-scale resistance screening using diverse *C. sublineola* strains would lead to the identification of resistant and susceptible genotypes. Integrated omics is an efficient approach to identify and characterize compatible (susceptibility) and incompatible (resistance) reaction of sorghum genotypes as well as to identify markers related to plant defense and pathogenicity factors. Availability of genomic information on sorghum and *C. sublineola* would enable efficient application of integrated omics approach to unravel their interactions. Validated markers for their association with the target trait can facilitate the identification of resistant genotypes in a diverse sorghum germplasm while defense related genes, particularly major *R*-genes are useful for genetic improvement of sorghum cultivars. However, the ever-evolving pathogen populations continue to pose a challenge to resistant cultivars deployed for production. Hence, it is important to add newly identified resistant cultivars to existing resistance gene pool and frequently re-evaluate their resistance against new *C. sublineola* strains to ensure durability of the resistance.

pathogens secrete effectors into the apoplast, considered the frontier of the interaction, investigating the host and pathogen proteins in this compartment would provide insights into early stages of the interaction process.

CONCLUSION AND FUTURE PERSPECTIVES

The sorghum–anthracnose interaction is economically important pathosystem. However, not much is known about the molecular aspects of the interaction processes between these organisms despite its significance in facilitating the improvement of the crop's resistance against the pathogen. The resistance of diverse sorghum landraces against anthracnose has been evaluated, and useful resistance sources for breeding were identified in several studies. Anthracnose resistance QTL regions are hotspots for potentially functional NB-LRR genes that, after functional validation, can be transferred, individually or by pyramiding multiple resistance genes, into elite sorghum materials. Moreover, the anthracnose defense-related genes identified so far can serve as biomarkers for quick and efficient identification of resistance sources in the diverse sorghum gene pool. The role of receptors, transcription factors, and phytohormones during compatible and incompatible interactions remains to be elucidated. Sorghum breeding efforts should consider

spatio-temporal distributions of the diverse *C. sublineola* populations in order to develop resilient cultivars with durable resistance and a wide adaptation. Existing genomic resources of both sorghum and *C. sublineola* is crucial to elucidate the interaction process using omics studies. Hence, the utilization of these resources has to be ramped up for their potential role in facilitating the identification of key components of host resistance and pathogen virulence, tracking dynamics of pathogen populations across seasons and geographical area, and accordingly contributing to accelerated and efficient sorghum breeding.

AUTHOR CONTRIBUTIONS

KA conceptualized and drafted the manuscript, with the assistance of MG. KA, MG, RO, and AC reviewed and edited the manuscript. KA prepared the figures and table. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Swedish Research Council grant (VR) grant #2019-05596 and partially by the SIDA for Research and Graduate Training Partnership between AAU and Swedish Universities (2018–2023).

REFERENCES

- Acharya, B., O'Quinn, T. N., Everman, W., and Mehl, H. L. (2019). Effectiveness of fungicides and their application timing for the management of sorghum foliar anthracnose in the mid-Atlantic United States. *Plant Dis.* 103, 2804–2811. doi: 10.1094/PDIS-10-18-1867-RE
- Afolayan, G., Deshpande, S. P., Aladele, S. E., Kolawole, A. O., Angarawai, I., Nwosu, D. J., et al. (2019). Genetic diversity assessment of sorghum (*Sorghum bicolor* (L.) Moench) accessions using single nucleotide polymorphism markers. *Plant Genet. Resour.* 17, 412–420. doi: 10.1017/S1479262119000212
- Ahn, E., Prom, L. K., Odvody, G., and Magill, C. (2019). Defense responses against the sorghum anthracnose pathogen in leaf blade and midrib tissue of johnsongrass and sorghum. *Physiol. Mol. Plant Pathol.* 106, 81–86. doi: 10.1016/j.pmpp.2018.12.008
- Ambawat, S., Sharma, P., Yadav, N. R., and Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiol. Mol. Biol. Plants* 19, 307–321. doi: 10.1007/s12298-013-0179-1
- Ananda, G. K. S., Myrans, H., Norton, S. L., Gleadow, R., Furtado, A., and Henry, R. J. (2020). Wild sorghum as a promising resource for crop improvement. *Front. Plant Sci.* 11:1108. doi: 10.3389/fpls.2020.01108
- Baillo, E. H., Hanif, M. S., Guo, Y. H., Zhang, Z. B., Xu, P., and Algam, S. A. (2020). Genome-wide identification of WRKY transcription factor family members in sorghum (*Sorghum bicolor* (L.) Moench). *PLoS One* 15:e0236651. doi: 10.1371/journal.pone.0236651
- Baroncelli, R., Sanz-Martin, J. M., Rech, G. E., Sukno, S. A., and Thon, M. R. (2014). Draft genome sequence of *Colletotrichum sublineola*, a destructive pathogen of cultivated sorghum. *Genome Announc.* 2, e00540–e00554. doi: 10.1128/genomeA.00540-14
- Basavaraju, P., Shetty, N. P., Shetty, H. S., de Neergaard, E., and Jorgensen, H. J. (2009). Infection biology and defence responses in sorghum against *Colletotrichum sublineolum*. *J. Appl. Microbiol.* 107, 404–415. doi: 10.1111/j.1365-2672.2009.04234.x
- Birker, D., Heidrich, K., Takahara, H., Narusaka, M., Deslandes, L., Narusaka, Y., et al. (2009). A locus conferring resistance to *Colletotrichum higginsianum* is shared by four geographically distinct *Arabidopsis* accessions. *Plant J.* 60, 602–613. doi: 10.1111/j.1365-3113.2009.03984.x
- Biruma, M., Martin, T., Fridborg, I., Okori, P., and Dixelius, C. (2012). Two loci in sorghum with NB-LRR encoding genes confer resistance to *Colletotrichum sublineolum*. *Theor. Appl. Genet.* 124, 1005–1015. doi: 10.1007/s00122-011-1764-8
- Boddu, J., Svabek, C., Ibraheem, F., Jones, A. D., and Chopra, S. (2005). Characterization of a deletion allele of a sorghum Myb gene, yellow seed1 showing loss of 3-deoxyflavonoids. *Plant Sci.* 169, 542–552. doi: 10.1016/j.plantsci.2005.05.007
- Boora, K. S., Frederiksen, R., and Magill, C. (1998). DNA-based markers for a recessive gene conferring anthracnose resistance in sorghum. *Crop Sci.* 38, 1708–1709. doi: 10.2135/cropsci1998.0011183X003800060048x
- Boyles, R. E., Brenton, Z. W., and Kresovich, S. (2019). Genetic and genomic resources of sorghum to connect genotype with phenotype in contrasting environments. *Plant J.* 97, 19–39. doi: 10.1111/tpj.14113
- Buiate, E. A. S., Souza, E. A., Vaillancourt, L., Resende, I., and Klink, U. P. (2010). Evaluation of resistance in sorghum genotypes to the causal agent of anthracnose. *Crop Breed. Appl. Biotechnol.* 10, 166–172. doi: 10.12702/1984-7033.v10n02a10
- Buiate, E. A. S., Xavier, K. V., Moore, N., Torres, M. F., Farman, M. L., Schardl, C. L., et al. (2017). A comparative genomic analysis of putative pathogenicity genes in the host-specific sibling species *Colletotrichum graminicola* and *Colletotrichum sublineola*. *BMC Genomics* 18:67. doi: 10.1186/s12864-016-3457-9
- Burrell, A. M., Sharma, A., Patil, N. Y., Collins, S. D., Anderson, W. F., Rooney, W. L., et al. (2015). Sequencing of an anthracnose-resistant sorghum genotype and mapping of a major QTL reveal strong candidate genes for anthracnose resistance. *Crop Sci.* 55, 790–799. doi: 10.2135/cropsci2014.06.0430
- Casela, C. R., and Frederiksen, R. A. (1993). Survival of *Colletotrichum graminicola* Sclerotia in sorghum stalk residues. *Plant Dis.* 77, 825–827. doi: 10.1094/Pd-77-0825
- Chala, A. (2013). Genetic diversity of *Colletotrichum sublineolum* isolates from a single field in southern Ethiopia and evidence for the existence of mat2 genotypes in different parts of the country. *Ethiop. J. Health Sci.* 36, 9–16.

- Chala, A., Tronsmo, A. M., and Brurberg, M. B. (2011). Genetic differentiation and gene flow in *Colletotrichum sublineolum* in Ethiopia, the Centre of origin and diversity of sorghum, as revealed by AFLP analysis. *Plant Pathol.* 60, 474–482. doi: 10.1111/j.1365-3059.2010.02389.x
- Chavan, S., Gray, J., and Smith, S. M. (2015). Diversity and evolution of Rp1 rust resistance genes in four maize lines. *Theor. Appl. Genet.* 128, 985–998. doi: 10.1007/s00122-015-2484-2
- Chopra, S., Gevens, A., Svabek, C., Wood, K. V., Peterson, T., and Nicholson, R. L. (2002). Excision of the Candystripe1 transposon from a hyper-mutable Y1-cs allele shows that the sorghumY1 gene controls the biosynthesis of both 3-deoxyanthocyanidin phytoalexins and phlobaphene pigments. *Physiol. Mol. Plant Pathol.* 60, 321–330. doi: 10.1016/S0885-5765(02)90411-X
- Crouch, J. A., and Beirn, L. A. (2009). Anthracnose of cereals and grasses. *Fungal Divers.* 39, 19–44.
- Crouch, J. A., and Tomaso-Peterson, M. (2012). Anthracnose disease of centipedegrass turf caused by *Colletotrichum eremochloae*, a new fungal species closely related to *Colletotrichum sublineola*. *Mycologia* 104, 1085–1096. doi: 10.3852/11-317
- Cuevas, H. E., and Prom, L. K. (2020). Evaluation of genetic diversity, agronomic traits, and anthracnose resistance in the NPGS Sudan Sorghum core collection. *BMC Genomics* 21:88. doi: 10.1186/s12864-020-6489-0
- Cuevas, H. E., Prom, L. K., Cooper, E. A., Knoll, J. E., and Ni, X. (2018). Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in the U.S. sorghum association panel. *Plant Genome* 11, 1–13. doi: 10.3835/plantgenome2017.11.0099
- Cuevas, H. E., Prom, L. K., and Cruet-Burgos, C. M. (2019). Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in NPGS Ethiopian sorghum germplasm. *G3* 9, 2879–2885. doi: 10.1534/g3.119.400350
- Cuevas, H. E., Prom, L. K., Erpelding, J. E., Brotons, V., and Lübberstedt, T. (2014). Assessments of genetic diversity and anthracnose disease response among Zimbabwe sorghum germplasm. *Plant Breed.* 133, 234–242. doi: 10.1111/pbr.12133
- Cuevas, H. E., Prom, L. K., Isakeit, T., and Radwan, G. (2016). Assessment of sorghum germplasm from Burkina Faso and South Africa to identify new sources of resistance to grain mold and anthracnose. *Crop Prot.* 79, 43–50. doi: 10.1016/j.cropro.2015.10.007
- Cuevas, H. E., Rosa-Valentin, G., Hayes, C. M., Rooney, W. L., and Hoffmann, L. (2017). Genomic characterization of a core set of the USDA-NPGS Ethiopian sorghum germplasm collection: implications for germplasm conservation, evaluation, and utilization in crop improvement. *BMC Genomics* 18:108. doi: 10.1186/s12864-016-3475-7
- da Costa, R. V., Zambolim, L., Cota, L. V., da Silva, D. D., Rodrigues, J. A. S., Tardin, F. D., et al. (2011). Genetic control of sorghum resistance to leaf anthracnose. *Plant Pathol.* 60, 1162–1168. doi: 10.1111/j.1365-3059.2011.02469.x
- Dao, T. T. H., Linthorst, H. J. M., and Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochem. Rev.* 10, 397–412. doi: 10.1007/s11101-011-9211-7
- Das, I. K., and Rajendrakumar, P. (2016). “Disease resistance in Sorghum” in *Biotic stress resistance in millets*. eds. I. K. Das and P. G. Padmaja (Academic Press), 23–67.
- Domazakis, E., Lin, X., Aguilera-Galvez, C., Wouters, D., Bijsterbosch, G., Wolters, P. J., et al. (2017). Effectoromics-based identification of cell surface receptors in potato. *Methods Mol. Biol.* 1578, 337–353. doi: 10.1007/978-1-4939-6859-6_29
- Erpelding, J. (2008). Sorghum germplasm resistance to anthracnose. *Am. J. Plant Sci. Biotechnol.* 2, 42–46.
- Evans, J., McCormick, R. F., Morishige, D., Olson, S. N., Weers, B., Hilley, J., et al. (2013). Extensive variation in the density and distribution of DNA polymorphism in sorghum genomes. *PLoS One* 8:e79192. doi: 10.1371/journal.pone.0079192
- FAOSTAT (2020). Food and Agriculture Organization of the United Nations. Available at: <http://www.fao.org/faostat/en/> (Accessed November 20, 2020).
- Felderhoff, T. J., McIntyre, L. M., Saballos, A., and Vermerris, W. (2016). Using genotyping by sequencing to map two novel anthracnose resistance loci in *Sorghum bicolor*. *G3* 6, 1935–1946. doi: 10.1534/g3.116.030510
- Fesel, P. H., and Zuccaro, A. (2016). Beta-glucan: crucial component of the fungal cell wall and elusive MAMP in plants. *Fungal Genet. Biol.* 90, 53–60. doi: 10.1016/j.fgb.2015.12.004
- Girma, G., Nida, H., Seyoum, A., Mekonen, M., Nega, A., Lule, D., et al. (2019). A large-scale genome-wide association analyses of Ethiopian sorghum landrace collection reveal loci associated with important traits. *Front. Plant Sci.* 10:691. doi: 10.3389/fpls.2019.00691
- Gong, A. D., Jing, Z. Y., Zhang, K., Tan, Q. Q., Wang, G. L., and Liu, W. D. (2020). Bioinformatic analysis and functional characterization of the CFEM proteins in maize anthracnose fungus *Colletotrichum graminicola*. *J. Integr. Agric.* 19, 541–550. doi: 10.1016/S2095-3119(19)62675-4
- Ibraheem, F., Gaffoor, I., and Chopra, S. (2010). Flavonoid phytoalexin-dependent resistance to anthracnose leaf blight requires a functional yellow seed1 in *Sorghum bicolor*. *Genetics* 184, 915–926. doi: 10.1534/genetics.109.111831
- Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Jupe, F., Chen, X. W., Verweij, W., Witek, K., Jones, J. D. G., and Hein, I. (2014). Genomic DNA library preparation for resistance gene enrichment and sequencing (RenSeq) in plants. *Methods Mol. Biol.* 1127, 291–303. doi: 10.1007/978-1-62703-986-4_22
- Kage, U., Kumar, A., Dhokane, D., Karre, S., and Kushalappa, A. C. (2016). Functional molecular markers for crop improvement. *Crit. Rev. Biotechnol.* 36, 917–930. doi: 10.3109/07388551.2015.1062743
- Lebeau, F. J., Coleman, O. H., and Agon, J. (1950). The inheritance of resistance in sorghum to leaf anthracnose. *CABI Direct* 42, 33–34.
- Li, T., Gonzalez, N., Inze, D., and Dubois, M. (2020). Emerging connections between small RNAs and phytohormones. *Trends Plant Sci.* 25, 912–929. doi: 10.1016/j.tplants.2020.04.004
- Li, L., Zhu, F. Y., Liu, H. J., Chu, A., and Lo, C. (2013). Isolation and expression analysis of defense-related genes in sorghum-*Colletotrichum sublineolum* interaction. *Physiol. Mol. Plant Pathol.* 84, 123–130. doi: 10.1016/j.pmp.2013.08.005
- Liab, J., Zhang, Q.-Y., Gaob, Z.-H., Wang, F., Duan, K., Ye, Z.-H., et al. (2013). Genome-wide identification and comparative expression analysis of NBS-LRR-encoding genes upon *Colletotrichum gloeosporioides* infection in two ecotypes of *Fragaria vesca*. *Gene* 527, 215–227. doi: 10.1016/j.gene.2013.06.008
- Liebrand, T. W. H., van den Burg, H. A., and Joosten, M. H. A. J. (2014). Two for all: receptor-associated kinases SOBIR1 and BAK1. *Trends Plant Sci.* 19, 123–132. doi: 10.1016/j.tplants.2013.10.003
- Little, C. R., Prom, L. K., Perumal, R., Tesso, T. T., Odvody, G. N., and Magill, C. W. (2012). Sorghum pathology and biotechnology—A fungal disease perspective: part I. grain mold, head smut, and ergot. *Eur. J. Plant Sci. Biotechnol.* 6, 31–44.
- Liu, W., Frick, M., Huel, R., Nykiforuk, C. L., Wang, X. M., Gaudet, D. A., et al. (2014). The stripe rust resistance gene Yr10 encodes an evolutionary-conserved and unique CC-NBS-LRR sequence in wheat. *Mol. Plant* 7, 1740–1755. doi: 10.1093/mp/ssu112
- Lo, S. C. C., De Verdier, K., and Nicholson, R. L. (1999). Accumulation of 3-deoxyanthocyanidin phytoalexins and resistance to *Colletotrichum sublineolum* in sorghum. *Physiol. Mol. Plant Pathol.* 55, 263–273. doi: 10.1006/pmp.1999.0231
- Mace, E. S., Tai, S. S., Gilding, E. K., Li, Y. H., Prentis, P. J., Bian, L. L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4:2320. doi: 10.1038/ncomms3320
- Mace, E., Tai, S. S., Innes, D., Godwin, I., Hu, W. S., Campbell, B., et al. (2014). The plasticity of NBS resistance genes in sorghum is driven by multiple evolutionary processes. *BMC Plant Biol.* 14:253. doi: 10.1186/s12870-014-0253-z
- Martin, T., Biruma, M., Fridborg, I., Okori, P., and Dixelius, C. (2011). A highly conserved NB-LRR encoding gene cluster effective against *Setosphaeria turcica* in sorghum. *BMC Plant Biol.* 11:151. doi: 10.1186/1471-2229-11-151
- McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., et al. (2018). The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* 93, 338–354. doi: 10.1111/tpj.13781
- Mehta, P. J., Wiltse, C. C., Rooney, W. L., Collins, S. D., Frederiksen, R. A., Hess, D. E., et al. (2005). Classification and inheritance of genetic resistance to anthracnose in sorghum. *Field Crop Res.* 93, 1–9. doi: 10.1016/j.fcr.2004.09.001
- Meng, X. Z., and Zhang, S. Q. (2013). MAPK cascades in plant disease resistance signaling. *Annu. Rev. Phytopathol.* 51, 245–266. doi: 10.1146/annurev-phyto-082712-102314

- Mengistu, G., Shimelis, H., Laing, M., and Lule, D. (2018). Breeding for anthracnose (*Colletotrichum sublineolum* Henn.) resistance in sorghum: challenges and opportunities. *Aust. J. Crop. Sci.* 12, 1911–1920. doi: 10.21475/ajcs.18.12.12.p1230
- Mengistu, G., Shimelis, H., Laing, M., and Lule, D. (2019). Assessment of sorghum genetic resources of Ethiopia for anthracnose (*Colletotrichum sublineolum* Henn.) resistance and agronomic traits. *J. Phytopathol.* 167, 667–678. doi: 10.1111/jph.12861
- Mengistu, G., Shimelis, H., Laing, M., Lule, D., and Mathew, I. (2020). Genetic variability among Ethiopian sorghum landrace accessions for major agromorphological traits and anthracnose resistance. *Euphytica* 216:113. doi: 10.1007/s10681-020-02650-6
- Mohan, S. M., Madhusudhana, R., Mathur, K., Chakravarthi, D. V. N., Rathore, S., Reddy, R. N., et al. (2010). Identification of quantitative trait loci associated with resistance to foliar diseases in sorghum [*Sorghum bicolor* (L.) Moench]. *Euphytica* 176, 199–211. doi: 10.1007/s10681-010-0224-x
- Moore, J. W., Dittmore, M., and TeBeest, D. O. (2008). Pathotypes of *Colletotrichum sublineolum* in Arkansas. *Plant Dis.* 92, 1415–1420. doi: 10.1094/PDIS-92-10-1415
- Mothaodi, T., Geleta, M., Chite, S., Fatih, M., Ortiz, R., and Bryngelsson, T. (2017). Genetic diversity in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from southern Africa as revealed by microsatellite markers and agromorphological traits. *Genet. Resour. Crop. Evol.* 64, 599–610. doi: 10.1007/s10722-016-0388-x
- Nagy, E. D., and Bennetzen, J. L. (2008). Pathogen corruption and site-directed recombination at a plant disease resistance gene cluster. *Genome Res.* 18, 1918–1923. doi: 10.1101/gr.078766.108
- Ordonio, R., Ito, Y., Morinaka, Y., Sazuka, T., and Matsuoka, M. (2016). Molecular breeding of *Sorghum bicolor*, A novel energy crop. *Int. Rev. Cell Mol. Biol.* 321, 221–257. doi: 10.1016/bs.ircmb.2015.09.001
- Panday, S., Sindhu, A., and Boora, K. (2002). RAPD based DNA markers linked to anthracnose disease resistance in *Sorghum bicolor* (L.) Moench. *Indian J. Exp. Biol.* 40, 206–211.
- Pandey, S. P., and Somssich, I. E. (2009). The role of WRKY transcription factors in plant immunity. *Plant Physiol.* 150, 1648–1655. doi: 10.1104/pp.109.138990
- Parra, L. B., Truco, M. J., Fletcher, K., and Michelmore, R. W. (2019). Identification of candidate genes for resistance to lettuce downy mildew using RenSeq k-mer association studies. *Mol. Plant-Microbe Interact.* 32:7.
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457, 551–556. doi: 10.1038/nature07723
- Patil, N. Y., Klein, R. R., Williams, C. L., Collins, S. D., Knoll, J. E., Burrell, A. M., et al. (2017). Quantitative trait loci associated with anthracnose resistance in sorghum. *Crop Sci.* 57, 877–890. doi: 10.2135/cropsci2016.09.0793
- Perumal, R., Menz, M. A., Mehta, P. J., Katile, S., Gutierrez-Rojas, L. A., Klein, R. R., et al. (2009). Molecular mapping of Cg1, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. *Euphytica* 165, 597–606. doi: 10.1007/s10681-008-9791-5
- Prom, L., Cuevas, H., Ramasamy, P., Thomas, I., and Clint, M. (2018). Inheritance of resistance of three sorghum lines to pathotypes of *Colletotrichum sublineola*, causal agent of anthracnose. *Plant Pathol. J.* 17, 75–79. doi: 10.3923/ppj.2018.75.79
- Prom, L., Erpelding, J., Perumal, R., Isakeit, T., and Cuevas, H. (2012a). Response of sorghum accessions from four African countries against *Colletotrichum sublineolum*, causal agent of sorghum anthracnose. *Am. J. Plant Sci.* 3, 125–129. doi: 10.4236/ajps.2012.31014
- Prom, L. K., Perumal, R., Erattimuthu, S. R., Little, C. R., No, E. G., Erpelding, J. E., et al. (2012b). Genetic diversity and pathotype determination of *Colletotrichum sublineolum* isolates causing anthracnose in sorghum. *Eur. J. Plant Pathol.* 133, 671–685. doi: 10.1007/s10658-012-9946-z
- Rosewich, U. L., Pettway, R. E., McDonald, B. A., Duncan, R. R., and Frederiksen, R. A. (1998). Genetic structure and temporal dynamics of a *Colletotrichum graminicola* population in a sorghum disease nursery. *Phytopathology* 88, 1087–1093. doi: 10.1094/Phyto.1998.88.10.1087
- Sherriff, C., Whelan, M. J., Arnold, G. M., and Bailey, J. A. (1995). rDNA sequence analysis confirms the distinction between *Colletotrichum graminicola* and *C. sublineolum*. *Mycol. Res.* 99, 475–178. doi: 10.1016/S0953-7562(09)80649-7
- Shi, H., Wang, L., Wang, Y., Cao, H., Hao, X., Zeng, J., et al. (2016). Transcriptome analysis of an anthracnose-resistant tea plant cultivar reveals genes associated with resistance to *Colletotrichum camelliae*. *PLoS One* 11:e0148535. doi: 10.1371/journal.pone.0148535
- Singh, M., Chaudhary, K., Singal, H. R., Magill, C. W., and Boora, K. S. (2006). Identification and characterization of RAPD and SCAR markers linked to anthracnose resistance gene in sorghum [*Sorghum bicolor* (L.) Moench]. *Euphytica* 149, 179–187. doi: 10.1007/s10681-005-9065-4
- Sperschneider, J., Dodds, P. N., Gardiner, D. M., Singh, K. B., and Taylor, J. M. (2018). Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol. Plant Pathol.* 19, 2094–2110. doi: 10.1111/mpp.12682
- Steuernagel, B., Witek, K., Krattinger, S. G., Ramirez-Gonzalez, R. H., Schoonbeek, H. J., Yu, G. T., et al. (2020). The NLR-annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiol.* 183, 468–482. doi: 10.1104/pp.19.01273
- Tesso, T., Perumal, R., Little, C. R., Adeyanju, A., Ghada, L. R., Prom, L., et al. (2012). Sorghum pathology and biotechnology—a fungal disease perspective: part II. Anthracnose, stalk rot, and downy mildew. *Eur. J. Plant Sci. Biotechnol.* 6, 31–34.
- Thakur, R., Rao, V., Wu, S., Subbara, K., Mathur, K., Tailor, H., et al. (2007). Genetic resistance to foliar anthracnose in sorghum and pathogenic variability in *Colletotrichum graminicola*. *Indian Phytopathol.* 60, 13–23.
- Tugizimana, F., Djami-Tchatchou, A. T., Steenkamp, P. A., Piater, L. A., and Dubery, I. A. (2018). Metabolomic analysis of defense-related reprogramming in *Sorghum bicolor* in response to *Colletotrichum sublineolum* infection reveals a functional metabolic web of phenylpropanoid and flavonoid pathways. *Front. Plant Sci.* 9:1840. doi: 10.3389/fpls.2018.01840
- Tugizimana, F., Steenkamp, P. A., Piater, L. A., Labuschagne, N., and Dubery, I. A. (2019). Unravelling the metabolic reconfiguration of the post-challenge primed state in *Sorghum bicolor* responding to *Colletotrichum sublineolum* infection. *Meta* 9:194. doi: 10.3390/metabo9100194
- Upadhyaya, H. D., Narsimha Reddy, K., Vetriventhan, M., Irshad Ahmed, M., Murali Krishna, G., Thimma Reddy, M., et al. (2017). Sorghum germplasm from west and Central Africa maintained in the ICRISAT genebank: status, gaps, and diversity. *Crop J.* 5, 518–532. doi: 10.1016/j.cj.2017.07.002
- Upadhyaya, H. D., Vetriventhan, M., and Deshpande, S. (2016). “Sorghum germplasm resources characterization and trait mapping” in *Sorghum Genome*. eds. S. Rakshit and Y. H. Wang (Cham: Springer), 77–94.
- Upadhyaya, H. D., Wang, Y. H., Sharma, R., and Sharma, S. (2013). Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. *Theor. Appl. Genet.* 126, 1649–1657. doi: 10.1007/s00122-013-2081-1
- Vaillancour, L., and Hanau, R. (1992). Genetic and morphological comparisons of *Glomerella* (*Colletotrichum*) isolates from maize and from sorghum. *Exp. Mycol.* 16, 219–222. doi: 10.1016/0147-5975(92)90030-U
- Valerio, H. M., Resende, M. A., Weikert-Oliveira, R. C. B., and Casela, C. R. (2005). Virulence and molecular diversity in *Colletotrichum graminicola* from Brazil. *Mycopathologia* 159, 449–459. doi: 10.1007/s1046-005-0373-y
- van der Biezen, E. A., and Jones, J. D. G. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* 23, 454–456. doi: 10.1016/S0968-0004(98)01311-5
- Verma, V., Ravindran, P., and Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC Plant Biol.* 16:86. doi: 10.1186/s12870-016-0771-y
- Wang, L., Chen, M., Zhu, F., Fan, T., Zhang, J., and Lo, C. (2020). Alternative splicing is a *Sorghum bicolor* defense response to fungal infection. *Planta* 251:14. doi: 10.1007/s00425-019-03309-w
- Wang, M. L., Dean, R., Erpelding, J., and Pederson, G. (2006). Molecular genetic evaluation of sorghum germplasm differing in response to fungal diseases: rust (*Puccinia purpurea*) and anthracnose (*Colletotrichum graminicola*). *Euphytica* 148, 319–330. doi: 10.1007/s10681-005-9040-0
- Wharton, P., and Julian, A. (1996). A cytological study of compatible and incompatible interactions between *Sorghum bicolor* and *Colletotrichum sublineolum*. *New Phytol.* 134, 25–34. doi: 10.1111/j.1469-8137.1996.tb01143.x
- Wharton, P., Julian, A., and O'Connell, R. (2001). Ultrastructure of the infection of *Sorghum bicolor* by *Colletotrichum sublineolum*. *Phytopathology* 91, 149–158. doi: 10.1094/PHYTO.2001.91.2.149
- Witek, K., Jupe, F., Witek, A. I., Baker, D., Clark, M. D., and Jones, J. D. G. (2016). Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol.* 34, 656–660. doi: 10.1038/nbt.3540
- Wu, J., Zhu, J. F., Wang, L. F., and Wang, S. M. (2017). Genome-wide association study identifies NBS-LRR-encoding genes related with anthracnose and

- common bacterial blight in the common bean. *Front. Plant Sci.* 8:1398. doi: 10.3389/fpls.2017.01398
- Xu, X. (2019). Identification and mapping of anthracnose resistance genes in sorghum [*Sorghum bicolor* (L.) Moench]. PhD thesis. Faculty of Purdue University.
- Xu, J., Qin, P., Jiang, Y., Hu, L., Liu, K., and Xu, X. (2020). Evaluation of sorghum germplasm resistance to anthracnose by *Colletotrichum sublineolum* in China. *Crop Prot.* 134:105173. doi: 10.1016/j.cropro.2020.105173
- Yang, S. H., Li, J., Zhang, X. H., Zhang, Q. J., Huang, J., Chen, J. Q., et al. (2013). Rapidly evolving R genes in diverse grass species confer resistance to rice blast disease. *Proc. Natl. Acad. Sci. U. S. A.* 110, 18572–18577. doi: 10.1073/pnas.1318211110
- Yazawa, T., Kawahigashi, H., Matsumoto, T., and Mizuno, H. (2013). Simultaneous transcriptome analysis of Sorghum and *Bipolaris sorghicola* by using RNA-seq in combination with de novo transcriptome assembly. *PLoS One* 8:e62460. doi: 10.1371/journal.pone.0062460
- Zheng, L. Y., Guo, X. S., He, B., Sun, L. J., Peng, Y., Dong, S. S., et al. (2011). Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). *Genome Biol.* 12:R114. doi: 10.1186/gb-2011-12-11-r114
- Zhu, F. Y., Li, L., Zhang, J. H., and Lo, C. (2015). Transgenic expression of a sorghum gene (SbLRR2) encoding a simple extracellular leucine-rich protein enhances resistance against necrotrophic pathogens in *Arabidopsis*. *Physiol. Mol. Plant Pathol.* 91, 31–37. doi: 10.1016/j.pmpp.2015.05.004
- Zipfel, C. (2014). Plant pattern-recognition receptors. *Trends Immunol.* 35, 345–351. doi: 10.1016/j.it.2014.05.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Abreha, Ortiz, Carlsson and Geleta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Nitrogen Use Efficiency in Sorghum: Exploring Native Variability for Traits Under Variable N-Regimes

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Shouvik Das,
Indian Agricultural Research Institute
(ICAR), India
Byoung-Cheorl Kang,
Seoul National University,
South Korea
Jaswinder Singh,
McGill University, Canada
Rajandeep Sekhon,
Clemson University, United States

*Correspondence:

Santosh P. Deshpande
s.deshpande@cgiar.org
Rajeev Gupta
g.rajeev@cgiar.org

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 17 December 2020

Accepted: 29 March 2021

Published: 21 April 2021

Citation:

Bollam S, Romana KK,
Rayaprolu L, Vemula A, Das RR,
Rathore A, Gandham P, Chander G,
Deshpande SP and Gupta R (2021)
Nitrogen Use Efficiency in Sorghum:
Exploring Native Variability for Traits
Under Variable N-Regimes.
Front. Plant Sci. 12:643192.
doi: 10.3389/fpls.2021.643192

Srikanth Bollam[†], Kirandeep Kaur Romana[†], Laavanya Rayaprolu, Anilkumar Vemula, Roma Rani Das, Abhishek Rathore, Prasad Gandham, Girish Chander, Santosh P. Deshpande* and Rajeev Gupta*

International Crops Research Institute for the Semi-Arid and Tropics, Patancheru, India

Exploring the natural genetic variability and its exploitation for improved Nitrogen Use Efficiency (NUE) in sorghum is one of the primary goals in the modern crop improvement programs. The integrated strategies include high-throughput phenotyping, next generation sequencing (NGS)-based genotyping technologies, and *a priori* selected candidate gene studies that help understand the detailed physiological and molecular mechanisms underpinning this complex trait. A set of sixty diverse sorghum genotypes was evaluated for different vegetative, reproductive, and yield traits related to NUE in the field (under three N regimes) for two seasons. Significant variations for different yield and related traits under 0 and 50% N confirmed the availability of native genetic variability in sorghum under low N regimes. Sorghum genotypes with distinct genetic background had interestingly similar NUE associated traits. The Genotyping-By-Sequencing based SNPs (>89 K) were used to study the population structure, and phylogenetic groupings identified three distinct groups. The information of grain N and stalk N content of the individuals covered on the phylogenetic groups indicated randomness in the distribution for adaptation under variable N regimes. This study identified promising sorghum genotypes with consistent performance under varying environments, with buffer capacity for yield under low N conditions. We also report better performing genotypes for varied production use—grain, stover, and dual-purpose sorghum having differential adaptation response to NUE traits. Expression profiling of NUE associated genes in shoot and root tissues of contrasting lines (PVK801 and HDW703) grown in varying N conditions revealed interesting outcomes. Root tissues of contrasting lines exhibited differential expression profiles for transporter genes [ammonium transporter (*SbAMT*), nitrate transporters (*SbNRT*); primary assimilatory (glutamine synthetase (*SbGS*), glutamate synthase (*SbGOGAT*[*NADH*], *SbGOGAT*[*Fd*]), assimilatory genes [nitrite reductase (*SbNiR*[*NADH*3]); and amino acid biosynthesis associated gene [glutamate dehydrogenase (*SbGDH*)]. Identification and expression

profiling of contrasting sorghum genotypes in varying N dosages will provide new information to understand the response of NUE genes toward adaptation to the differential N regimes in sorghum. High NUE genotypes identified from this study could be potential candidates for in-depth molecular analysis and contribute toward the development of N efficient sorghum cultivars.

Keywords: sorghum, genetic variability, N content, NUE, expression analysis

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is one of the important staple food crops and fifth-most cultivated cereal after wheat, rice, maize, and barley (Taylor et al., 2006). Sorghum is a multi-purpose commodity in terms of its utility, such as grain sorghum for human food and animal feed, forage sorghum for forage and fodder, sweet stalk sorghum providing fiber, and feedstock for biofuel. Genetic and phenotypic variability in sorghum is evident by its spread from the North America to the African continent, through the Middle East, the Indian sub-continent, and further parts of East Asia to Australia resulted in distinct botanical races (Kimber, 2000; Morris et al., 2013). Sorghum follows C_4 type photosynthesis pathway, and its efficient use of nutrients, radiation, and water makes it adaptable to harsh and water-limited conditions (Paterson et al., 2008). Due to cultivated sorghum's small genome size (812 Mbp) and diploid nature ($2n = 20$), sorghum is used as a model for genome analysis. The deep root system architecture of sorghum makes it drought-tolerant and adaptable to grow in a water-limited environment. Despite its C_4 nature and relatively better drought tolerance compared to maize (Paterson et al., 2009), sorghum still depends mainly on nitrogen (N) fertilizer for achieving higher grain yields in an intensive agricultural system. Nitrogen is an essential macronutrient, most abundantly absorbed by roots, and 75% of the N present in the leaf is allocated to chloroplasts. N is the primary constituent of most of the important biomolecules viz., nucleotides, amino acids, proteins, and hormones related to the plants overall growth and development. About 1.5–2.0% of total plant dry matter and 16% of the plant protein was covered by N (Frink et al., 1999).

In the last four decades, breeding efforts, along with the use of synthetic N fertilizers, enabled substantial increment in crop productivity, especially in irrigated production systems, contributing to the “Green Revolution” that addressed the global food needs. However, 50–70% of applied nitrogen fertilizer lost to the environment through volatilization, leaching, groundwater runoff, and nitrous oxide emissions from N fertilizer residues, in-turn, pose adverse effects on the environment (Zhang et al., 2013). High fertilizer N application is one of the major input costs to farmers, and it also affects soil health by acidification. With the priority of lower N fertilizer input and environmentally friendly agriculture (in the intensive production systems), the development of crops and/or genotypes with high NUE (especially in the subsistence farming) and better yield is critical for the sustainable production of sorghum across diverse agroecosystems in the world. NUE in crop plants is

a very complex phenomenon, governed by the economic produce of the species (grain, forage/fodder, or dual-purpose) and is defined as the quantity of biomass and/or grain produced per unit of available N in the soil (Moll et al., 1982; Good et al., 2004). Crop response to N mainly depends on the genotype and its interaction with applied N fertilizer (Masclaux-Daubresse and Chardon, 2011). Under the high N conditions, most commonly observed in the intensive agricultural systems, variation in NUE is primarily attributed to differences in N uptake capacity. In contrast, under limited N conditions, prevalent in sorghum production areas in Asia and Africa, NUE variation is driven by changes in N remobilization and utilization efficiency. There is a huge genetic variability is present for NUE, and associated traits in cereals [sorghum (Youngquist et al., 1992), rice (Tirol-Padre et al., 1996; Rao et al., 2018), wheat (Le Gouis et al., 2000), and maize (McCullough et al., 1994)]. This genetic variability is a valuable source to help understand physiological, molecular, and genetic basis of NUE and its further exploitation for the development of high NUE crops.

NUE is a complex trait and is driven by many genes associated with N uptake, assimilation, and remobilization. A comprehensive understanding of physiological and molecular mechanisms under pinning N stress tolerance and/or NUE in sorghum is critical for its effective exploitation, facilitated through available genetic and genomic resources. By utilizing the next generation sequencing (NGS) technologies and gene identification strategies, putative genes associated with NUE in sorghum were identified and characterized (Gelli et al., 2014, 2016, 2017; Massel et al., 2016; Diatloff et al., 2017). However, very few studies with expression studies at varying N conditions in different tissue samples are conducted. In the present study, we have conducted a series of field and lab experiments to evaluate the genotype, nitrogen treatment (0, 50, and 100% of the recommended N) and season (two different seasons) specific variations associated with NUE in sixty diverse sorghum accessions. Wide variability was observed for physiological, agronomical, growth, and biological yield parameters associated with NUE in sorghum. We have generated genotyping by sequencing (GBS) based single nucleotide polymorphism (SNP) data to study the genetic diversity for this panel of sorghum genotypes. High and low NUE genotypes were identified based on their performance under low N conditions. We have also identified grain, fodder and dual-purpose sorghum lines which could be potential donors for crop improvement programs. We also studied the expression profiles of NUE associated genes in shoot and root samples of contrasting sorghum genotypes under varying N conditions.

MATERIALS AND METHODS

Plant Material

The plant material included a diverse set of 60 sorghum accessions (**Supplementary Table 1**). This set included parents of mapping populations such as Back-Cross derived Nested Association Mapping (BCNAM) populations; bi-parental mapping populations, and accessions from different countries such as India, Niger, Sudan, South Africa, Pakistan, Yemen, Cameroon, Nigeria, United States, Lesotho, Ethiopia, Mali, and the United States.

Experimental Design

The set of 60 sorghum accessions were field evaluated in a split-plot alpha lattice design under three N fertilizer application levels (0, 50, and 100% of the recommended (90 kg ha^{-1}) N) with three replications for two seasons (2016–17 and 2017–18) in the black soil precision fields of International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India. Before starting the experiment, soil testing was done by collecting samples from 0 to 12 inches, and 12–24 inches, across the field as per the standard sampling procedure. The results were provided in **Supplementary Table 2**. An individual evaluation test plot with 2 m length and four-rows with 0.60 m inter-row spacing were sown at a density of 15–20 seeds (with 0.15 m plant to plant distance) per row for each accession. All entries were tractor planted on the same day in tilled plots. Other fertilizers such as phosphorus (P) and potassium (K) were applied to all plots at the rate of 50 and 40 kg ha^{-1} , respectively (as per the recommendation of PJ Telangana State Agricultural University, Rajendra Nagar, Hyderabad for the trial production ecology). The basal application of P and K fertilizers except N was applied in 2 equal splits at 20 days from the date of emergence and 1 month after application of the first N dose. The sources of N, P, and K were urea (46% N), single superphosphate (16% P_2O_5), and murate of potash (60% K_2O). Weeding was done at 20–30 days interval. During 2016–17, sowing was done on 15th December 2016 and the trial was harvested on 27th April 2017. Similarly, during 2017–18, sowing took place on 17th November 2017 and the trial was harvested on 4th April 2018. Irrigation was given four times in each season during 2016–17 (16th December 2016, 29th December 2016, 18th January 2017, and 02nd February 2017) and 2017–18 (18th November 2017, 04th December 2017, 20th December 2017, 05th January 2018) (**Supplementary Figures 1, 2**).

Phenotyping of the Traits Associated With NUE

Different physiological, agronomical, and biological yield attributes associated with NUE were systematically recorded in three different N dosages (0, 50, and 100% of the recommended N) for two seasons¹.

Leaf Parameters

Chlorophyll content (CC) was recorded on the flag leaf of three random plants of the middle two rows of the plot using SPAD meter (Konica Minolta Sensing Americas, Inc., Ramsey, NJ) at the anthesis stage (around 75 days after emergence). Leaf area (LA); specific leaf area for each genotype was measured by running all the harvested leaves of a plant through benchtop leaf area meter Li3100C (LI-COR Inc., Lincoln, NE, United States). Leaf number (LN); where leaf number of each genotype harvested for LA were counted.

Growth Parameters

During flowering stage, days to 50% flowering (DFL); number of days from the emergence date to the day on which 50% of the plants in a plot reached anthesis at least halfway. At the harvest stage, Plant height (PH) was measured from the base of the plant to the tip of the main head as an average of three plants, randomly chosen from the center two rows of each plot and expressed in cm. Plant stand (PS); the total number of plants in the center two rows of each plot was recorded. Number of tillers (NT); the total number of tillers in the center two rows of each plot were recorded.

Panicle Parameters

Panicle number (PN); the total number of heads in the center two rows of each plot were recorded. Panicle weight (PW); weight (g/plot) of all the panicles in the center two rows of each plot were measured.

Biological Yield Parameters

Fresh straw yield (FSY); weight (g/plot) of freshly harvested straw was recorded for center two rows of each plot. Dry straw yield (DSY); weight (g/plot) of 10 days sundried straw was recorded for center two rows of each plot. Grain yield (GY); grain weight (g/plot) of all the panicles which are sun-dried and threshed from the center two rows of each plot. Test weight (TW); 200 seed weight (g) of each genotype was recorded. Harvest Index [HI (%)] is the ratio of grain yield to the total biomass, considered as the measure of biological success.

NUE Parameters

N content in grain (GN%) and N content in straw (SN%) of each genotype was estimated by sulfuric acid-selenium digestion method (Sahrawat et al., 2002). Grain and straw samples were fine powdered using clone mixture (Cyclone sample mill), 250 mg of fine powder was used for N estimation at Charles Renard Analytical laboratory, ICRISAT. The samples were digested with sulfuric acid-selenium and then analyzed using an Auto-analyzer (Skalar SAN System, AA Breda, Netherlands).

Genotyping-by-Sequencing and Single Nucleotide Polymorphism Identification

DNA was isolated from leaves of each accession at 4–6 leaf stage using the modified hexadecyltrimethyl ammonium bromide (CTAB) protocol (Mace et al., 2003). Genotyping was performed by following the GBS approach (Elshire et al., 2011), restriction enzyme *ApeKI* (NEB R0643L) used for complexity

¹<https://doi.org/10.21421/D2/WCAHRK>

reduction. The GBS library was sequenced on IlluminaHiSeq 2500 (Illumina Inc., San Diego, CA, United States) following the manufacturer's protocol. SNPs were called using the TASSEL v5.2 GBS pipeline (Bradbury et al., 2007) against sorghum assembly v3.1 (Phytozome.JGI.doe.gov. Phytozome, 2019). The final data included a total of 89,770 SNPs (obtained from 58 accessions out of 60) with minor allele frequency (MAF) > 1%, and missing data <50% were used in this study.

Diversity and Population Structure

An unweighted neighbor-joining phylogenetic tree was constructed in TASSEL v5.2 (TASSEL v5.2). The hierarchical population structure was estimated by using the ADMIXTURE program, a model-based estimation of ancestry in unrelated individuals using the maximum-likelihood method (Alexander et al., 2009). ADMIXTURE implements a cross-validation (CV) feature, together with the number of iterations to convergence, allowing to determine the number of subpopulations (k -values) that best fits the data. The Admixture analysis was performed for different K (number of sub-populations) varying from 2 to 8. The most appropriate K -value was selected after considering 10-fold cross-validations whereby the best K exhibits low cross-validation error compared to other K -values and good correspondence with the clustering pattern obtained by hierarchical cluster tree.

Contrasting Genotypes Screening for N Stress Under Hydroponics System and Tissue Sample Collection for Expression Studies

Based on grain yield data under low N conditions, two contrasting genotypes for NUE (High NUE/better performer:

PVK801 and low NUE/poor performer: HDW703) were selected for expression profiling. Sorghum accessions were germinated on the sand. Eight-days-old seedlings with uniform length (both plumule and radicle) were selected and transferred to the nutrient solution (Modified Hoagland) in the glasshouse. The seedlings were maintained under a 16/8 h photo-period cycle at 25°C (day) and 18°C (night). The pH of the nutrient solution adjusted to 5.8 and refreshed every 3 days. Two weeks old seedlings were transferred to a modified Hoagland solution with 0 and 100% of recommended N conditions. The plants grown in 0% N generated N stress symptoms. From 24 days old seedlings (**Figure 1**), shoot and root samples were collected separately and frozen in liquid nitrogen and stored at -80°C until RNA isolation.

Total RNA Isolation, cDNA Synthesis, and Primer Design

Total RNA was isolated from the shoot and root tissues using the RN easy Plant Mini kit (Qiagen, Germany). The quality of the isolated RNA was determined by Nanodrop® ND100 spectrophotometer (Thermo Fisher Scientific, United States) and RNA was treated with RNase free DNase enzyme (Thermo Fisher Scientific, United States). One microgram of total RNA was used for first-strand cDNA synthesis using the Superscript III RT kit (Life Technologies, United States) as per the manufacturer's instructions. A total of 13 genes associated with N uptake, assimilation, and remobilization were selected *a priori*, based on previously published studies. Gene-specific primers were designed (**Supplementary Table 3**) using Primer3 software² with the specifications such as—G + C content 45–55%, Amplicon size 80–170 base pairs (bp),

²<http://simgene.com/Primer3>

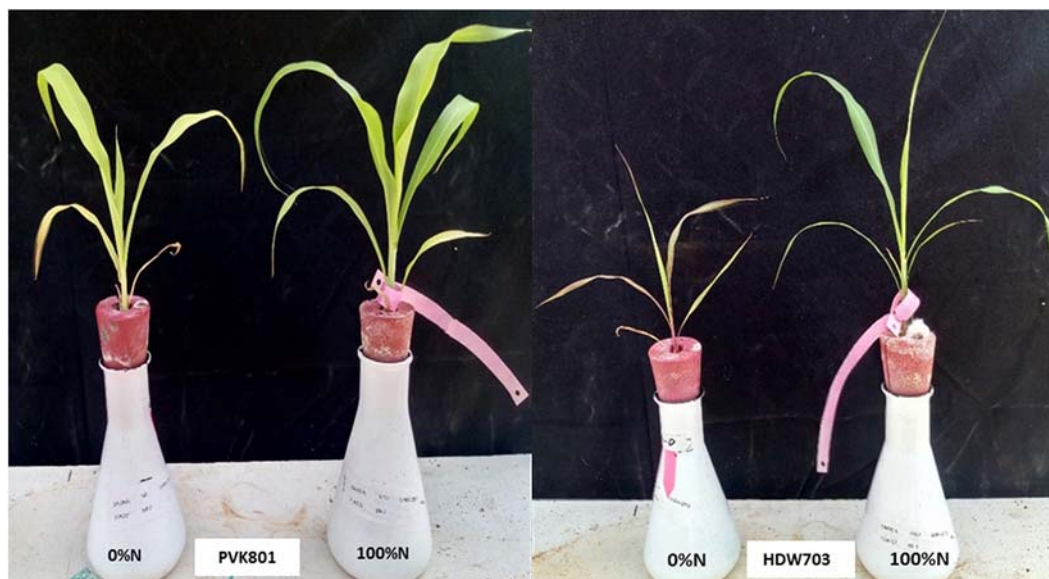


FIGURE 1 | High NUE/better (PVK801) and low NUE/poor (HDW703) performers grown under varying N (N0 and N100) conditions in hydroponics system using modified Hoagland solution.

length of primer 19–24 nucleotides and melting temperature (T_m) 58–62°C.

Quantitative Real-Time PCR (qRT-PCR) Analysis

All the qRT-PCR reactions were performed on Applied Biosystems 7500 Real-Time PCR (Life Technologies, United States), in 96 well optical plates. Reactions were performed in a final volume of 20 μ l, containing 60 ng of cDNA samples, 10 μ l of SYBR® Green qPCR SuperMix (Invitrogen, United States), 500 nM of each primer. The qRT-PCR cycling conditions included: 50°C for 2 min (Pre-incubation), 95°C for 10 min (denaturation) followed by 40 cycles of 95°C for 15 s (denaturation), 60°C for 1 min (annealing and extension). For all the samples, qRT-PCR was performed on three biological replicates and three technical replicates. Controls were set up for each sample in duplicates using the *SbUbiquitin* gene (internal control) for normalizing the gene expressions. qRT-PCR data were analyzed by 7500 Sequence Detection Software (Applied Biosystems, United States) with default baseline and threshold, relative expression of genes calculated by the $2^{-\Delta\Delta CT}$ method (Schmittgen and Livak, 2008).

Statistical Analysis

Analysis of variance (ANOVA) for all the traits was performed for individual and across seasons using PROC MIXED procedure of SAS version 9.4 (SAS Institute Inc [SAS], 2018), considering the season, treatments (whole plot), genotype (subplot), replications, as fixed effects and block as a random effect. Individual season variances were modeled into the combined analysis. Best Linear Unbiased Estimates (BLUE's) were calculated for the main and interaction effects of season, treatment, and genotype. Multiple comparisons were performed for significant effects ($p < 0.05$). Correlation coefficient analysis was performed using the PROC CORR procedure.

RESULTS

Evaluation of Sorghum Genotypes in Field Conditions for Two Seasons

Effects of Nitrogen and Genotypes on Leaf Parameters

The post-flowering SPAD showed an increase in mean and range values among N application rates from 0 to 100 in both individual seasons and pooled data with 11–16% variation (Table 1). The ANOVA revealed highly significant genotype \times treatment interaction for the season 2017 and pooled data. A gradual increase in leaf area was observed in 2017 and across seasons, but a slight decrease was observed for the N50 treatment in 2016. The genotype \times treatment was significant in individual and pooled seasons along with significant interaction effects for season \times genotype and season \times genotype \times treatment. An exponential increase in leaf number with N dosage in individual seasons and

across seasons with significant variance for season \times genotype (Table 2) was observed.

Effects of Nitrogen and Genotypes on Growth Parameters

The flowering time exhibited a minimum variation with N treatments in both individual and pooled seasons, with genotype \times treatment interaction being non-significant ($p \geq 0.05$) except during the 2017 trial. A gradual increase in plant height was observed with increasing dosage of N with significant genotype \times treatment, season \times genotype, and season \times genotype \times treatment interactions in pooled seasons. The number of tillers increased with a higher quantity of N in both individual and pooled seasons, and significant genotype \times treatment interaction was noted for 2017 (Table 2).

Effects of Nitrogen and Genotypes on Panicle Parameters

With increase in dosage of nitrogen, panicle number also increased in individual and across seasons with significant season \times genotype and season \times genotype \times treatment interactions. Panicle number observed more than 15% increase at N50 compared to N0 in individual seasons (Table 1). Accordingly, in both individual and across seasons, the panicle weight also increased with higher N concentrations. Significant genotype \times treatment, season \times genotype and season \times genotype \times treatment interactions were noted (Table 2).

Effects of Nitrogen and Genotypes on Biological Yield Parameters

Both grain yield and test weight exhibited an increasing trend with higher quantities of N dosages with 15–20% variation for grain yield. Significant interactions for genotype \times treatment, season \times genotype and season \times genotype \times treatment interactions were observed except in 2016 for grain yield.

In individual seasons fresh biomass yield recorded gradual increase with higher N-dosages, but the magnitude of increase from N0 to N50 was more compared to N50 to N100 dosages. The genotype \times treatment, season \times genotype and season \times genotype \times treatment interactions were significant overall except for 2016. The same trend was observed for dry stover yield where a gradual increase in weight was observed in individual seasons. A wide range of 11–24% variation was observed between N0 to N50 treatments in all seasons (Table 1). Harvest index had a steady increase with increasing dosages of nitrogen with season \times genotype and season \times genotype \times treatment significant (Table 2).

Effects of Nitrogen and Genotypes on NUE Parameters

The grain N content plunged from N0 to N50, while it increased from N50 to N100 in individual and across seasons (Table 1). The stalk N content was similar in range for N0 and N50 across seasons, whereas in 2017, there was a decrease from N0 to N50 and a rise to N100 treatments. All the interaction

TABLE 1 | Summary performance of 60 sorghum genotypes grown in different N conditions (0 N, 50% N, and 100% of the recommended N) for the season 2016–17, 2017–18, and across seasons.

Trait	Treatment	2016–17		2017–18		Across seasons	
		Range	Mean	Range	Mean	Range	Mean
SPAD (at anthesis stage)	N0	29.23–47.10	38.54	16.73–39.86	28.33	26.12–41.30	33.44
	N50	34.55–50.05	43.13	23.58–47.71	33.32	31.91–46.55	38.23
	N100	41.54–55.15	48.89	27.47–56.42	40.85	36.84–53.87	44.88
Leaf Area (cm ²)	N0	643.50–6459.72	2055.33	599.43–3187.43	1644.83	706.13–4486.14	1850.08
	N50	587.28–3408.67	1881.42	669.93–4212.95	2082.35	944.93–3431.58	1981.89
	N100	1126.86–8967.24	2444.70	928.84–5479.43	2345.36	1221.04–6082.23	2395.03
Leaf count (no.)	N0	6.35–26.93	12.78	7.13–21.90	12.10	8.23–23.13	12.44
	N50	6.48–21.22	11.52	7.00–26.00	13.58	7.49–21.68	12.53
	N100	7.82–26.21	12.72	8.26–39.83	16.60	8.88–25.11	14.60
Number of panicles	N0	16.17–131.27	29.059	16.30–198.43	37.31	19.02–107.00	33.07
	N50	17.10–52.97	31.51	17.36–161.73	42.21	19.26–99.95	36.67
	N100	21.98–126.78	35.34	19.63–91.81	43.35	21.94–108.59	39.24
Wt. of panicles (g)	N0	216.97–1847.38	887.77	174.38–1347.77	590.93	382.68–1325.59	739.35
	N50	282.29–1820.03	998	265.39–1827.33	847.45	372.91–1657.17	922.89
	N100	617.00–2323.72	1393.78	371.06–1951.89	924.23	600.47–1734.42	1159.00
Grain yield (g)	N0	125.54–1460.42	641.09	13.42–1159.42	359.40	195.14–972.92	500.24
	N50	179.35–1382.67	742.36	48.11–1244.52	555.09	169.73–1218.68	649.17
	N100	245.58–1879.09	1061.09	25.23–1355.34	556.47	147.42–1569.09	809.19
Test weight (g)	N0	0.78–3.98	2.56	0.90–4.15	2.54	0.84–4.03	2.55
	N50	1.03–3.98	2.64	1.29–4.43	2.72	1.16–4.14	2.68
	N100	1.02–4.01	2.74	1.20–4.64	2.73	1.18–4.31	2.74
Fresh stalk yield (kg)	N0	2946.28–9832.02	5442.16	1873.04–6797.77	4349.18	2842.85–7850.58	4895.67
	N50	3876.03–11764.00	6387.30	2626.82–10838.00	5677.98	3379.89–11205	6032.62
	N100	4223.51–16276.00	7808.41	2175.64–10654.00	5819.94	3199.58–12943.00	6814.17
Dry stalk yield (kg)	N0	1481.06–4522.73	2980.12	1071.46–3938.98	2273.48	1583.76–4038.23	2626.80
	N50	2169.29–6394.11	3348.63	1267.65–6009.78	2858.36	1755.07–5598.91	3103.50
	N100	2284.02–6476.93	3878.90	1207.92–5080.20	2956.85	1916.44–5712.39	3417.88
Harvest index	N0	3.87–42.75	19.99	1.05–50.45	15.38	5.84–35.86	17.68
	N50	6.73–39.48	21.48	0.10–60.99	19.70	5.40–40.85	20.95
	N100	7.60–50.89	26.61	0.26–63.21	20.42	6.85–49.16	23.16
N content in grain	N0	1.17–2.12	1.63	1.05–2.25	1.37	1.20–2.09	1.50
	N50	1.16–2.24	1.65	0.90–1.93	1.35	1.04–1.89	1.50
	N100	1.18–2.18	1.64	1.13–1.90	1.46	1.23–1.90	1.55
N content in stalk	N0	0.40–1.36	0.72	0.19–0.64	0.38	0.32–0.96	0.55
	N50	0.38–1.25	0.76	0.17–0.72	0.34	0.31–0.90	0.55
	N100	0.38–1.18	0.73	0.21–0.94	0.45	0.32–0.90	0.59

Data presented here is the mean of the three replications.

effects were significant, except in 2016 for stalk N. Genotype \times treatment interactions are significant in 2017 for grain N content (Table 2).

Correlation Coefficient Analysis

The correlation coefficient analysis with three different N treatments revealed a significant positive correlation between leaf area and flowering time in individual and across seasons. Under N0, N50, and N100 treatments, fresh and dry biomass yield had significant correlations with flowering time and leaf area except for 2017 under N100 treatment ($p \geq 0.05$). Plant height was also positively correlated with biomass yields except in the 2017 trial under N0 and N100 conditions. In the 2017 season, biomass

yields were negatively correlated with grain yield and panicle weight except in N50 dosages ($p \geq 0.05$). Under all treatments, fresh biomass had a significant positive correlation with dry biomass yield in individual and across seasons. A similar trend was observed for panicle weight with grain yield; and the number of tillers with panicle number and leaf number. Grain N percent had significant negative correlations to panicle weight and grain yield (except for N0 treatment for grain yield). Stalk N content was positively correlated with grain N content and had a negative correlation with plant height (except for N100 with $\text{prob} \geq 0.05$). A significant positive correlation was observed between stalk N content and panicle number under N0 treatment in 2017 and across seasons (Tables 3–5 and Supplementary Tables 4, 5).

TABLE 2 | Summary of ANOVA for leaf, growth, panicle, biological yield parameters along with NUE traits under three N dosages in field for the season 2016–17, 2017–18, and across seasons.

S. No.	Trait	2016–17		2017–18		Across seasons						
		Genotype (G)	G * T	Genotype (G)	G * T	Genotype (G)	Treatment (T)	Season (S)	G * T	S * G	S * T	S * G * T
Leaf parameters												
1	SPAD (at anthesis stage)	2.16	0.81 ^{ns}	10.99	5.81	7.03	51.81	14.26 ^{ns}	2.48	4.1	0.53	2.96
2	Leaf Area (cm²)	10.71	2.92	11.27	3.75	15.74	37.23	0.16 ^{ns}	3.19	5.91	11.14	3.54
3	Leaf count (no.)	3.13	0.97 ^{ns}	5.02	1.68 ^{ns}	5.19	2.60 ^{ns}	1.13	1.27 ^{ns}	2.26	2.22	1.25 ^{ns}
Growth parameters												
1	Days to 50% flowering (days)	23.22	0.92 ^{ns}	37.27	1.42	36.34	0.31 ^{ns}	0.45 ^{ns}	1.15 ^{ns}	6.49	1.83	1.03 ^{ns}
2	Plant height (cm)	76.86	1.22 ^{ns}	176.9	13.8	192.18	2.35 ^{ns}	1329.51	6.61	12.38	0.89	6.37
3	Number of tillers	7.87	1.02 ^{ns}	17.95	1.86	21.6	14.98 ^{ns}	3847.82	1.63 ^{ns}	5.01	1.28	1.53 ^{ns}
Panicle parameters												
1	Number of panicles	9.2	1.54 ^{ns}	0.97 ^{ns}	1.01 ^{ns}	8.79	20.95 ^{ns}	163.30 ^{ns}	1.04 ^{ns}	2.04	1.21	1.26
2	Wt. of panicles (g)	4.52	1.53 ^{ns}	26.75	4.08	13.42	249.81	14.02	3.58	8.6	36.51	3.43
Biological yield parameters												
1	Grain yield (g)	4.81	1.52 ^{ns}	40.31	5.66	23.01	118.46	23.09 ^{ns}	3.41	11.25	34.05	4.28
2	Test weight (g)	6.23	3.82	31.36	5.47	82.23	38.7	0.24 ^{ns}	4.33	11.78	3.71	3.62
3	Fresh stalk yield (kg)	10.02	1.47 ^{ns}	27.24	3.28	24.1	39.77	8.58 ^{ns}	2.48	6.04	4.61	1.85
4	Dry stover yield (kg)	11.41	1.08 ^{ns}	22.66	2.49	25.67	29.43 ^{ns}	536.84 ^{ns}	1.92	4.5	2.16	1.62 ^{ns}
5	Harvest Index	8	1.09 ^{ns}	29.81	2.43	24.64	12.99 ^{ns}	2.14	1.66 ^{ns}	12.39	3.75	2
NUE traits												
1	N content in grain	8.63	1.12 ^{ns}	11.29	1.76	9.48	1.01 ^{ns}	3.89 ^{ns}	1.52 ^{ns}	9.73	0.99	1.19 ^{ns}
2	N content in stalk	2.75	1.5 ^{ns}	13.04	5.32	5.58	6.91 ^{ns}	30515 ^{ns}	2.99	3.64	14.45	2.5

"ns" denotes for not significant. All Fstat/genetic variance are significant @p of 0.001 except for traits with "ns."

TABLE 3 | Correlation coefficient analysis of 60 genotypes for different traits at N0 condition across seasons.

Trait	DFL	CC	LA	PHT	PN	GY	PW	DSY	FSY	LN	NT	HI	TW	GN%	SN%
DFL	1														
CC	0.16	1													
LA	0.72**	0.28*	1												
PHT	0.01	-0.23	-0.05	1											
PN	-0.28*	-0.11	-0.34**	-0.05	1										
GY	-0.29*	-0.05	-0.29*	-0.08	-0.02	1									
PW	-0.09	0.05	-0.28*	0.04	0.04	0.68**	1								
DSY	0.56**	0.06	0.43**	0.48**	-0.27*	-0.28*	-0.13	1							
FSY	0.59**	0.12	0.38**	0.42**	-0.21	-0.3*	-0.14	0.92**	1						
LN	0.27*	0.07	0.36**	0.07	0.34**	-0.35**	-0.34**	0.16	0.19	1					
NT	-0.1	0.03	-0.11	-0.04	0.52**	-0.16	-0.06	-0.15	-0.05	0.52**	1				
HI	-0.38**	-0.16	-0.41**	-0.17	-0.02	0.87**	0.62**	-0.58**	-0.55**	-0.43**	-0.09	1			
TW	-0.12	0.14	0.05	-0.08	-0.38**	0.2	0.19	-0.06	-0.15	-0.2	-0.29*	0.22	1		
GN%	0.21	0.18	0.28*	-0.04	-0.18	-	-0.14	0.17	0.1	0.06	0.02	-0.27*	0.03	1	
SN%	0.08	-0.07	0.13	-0.28*	0.39**	-0.05	-0.18	-0.21	-0.2	0.18	0.24	-0.02	-0.07	0.26*	1

DFL, Days to 50% flowering; CC, SPAD at anthesis stage; LA, Leaf area; PHT, Plant height; PN, Panicle number; GY, Grain yield (g); PW, Panicle weight (g); DSY, Dry stalk yield (g); FSY, Fresh stalk yield (g); LN, Leaf number; NT, Number of tillers; HI, Harvest index; TW, Test weight (100 seed weight); GN%, N content in grain, SN%, N content in stalk. * denotes $p = 0.05$, ** denotes $p = 0.001$.

TABLE 4 | Correlation coefficient analysis of 60 genotypes for different trait at N50 condition across seasons.

Trait	DFL	CC	LA	PHT	PN	GY	PW	DSY	FSY	LN	NT	HI	TW	GN%	SN%
DFL	1														
CC	-0.05	1													
LA	0.65**	0.02	1												
PHT	0.02	-0.14	0.08	1											
PN	-0.29*	-0.18	-0.36**	-0.01	1										
GY	-0.09	0.02	-0.02	-0.23	-0.09	1									
PW	0.003	-0.06	0.05	-0.23	-0.16	0.84**	1								
DSY	0.43**	-0.05	0.47**	0.5**	-0.13	-0.25	-0.15	1							
FSY	0.41**	-0.0004	0.38**	0.44**	-0.03	-0.3*	-0.2	0.93**	1						
LN	0.16	-0.09	0.29*	-0.001	0.39**	-0.12	-0.2	0.17	0.17	1					
NT	-0.11	-0.13	-0.11	-0.01	0.67**	-0.06	-0.1	0.09	0.17	0.58**	1				
HI	-0.24	-0.06	-0.25	-0.31*	-0.08	0.85**	0.71**	-0.6**	-0.62**	-0.35**	-0.18	1			
TW	-0.19	0.15	0.1	0.04	-0.47**	0.16	0.29*	0.01	-0.11	-0.16	-0.3*	0.11	1		
GN%	0.11	0.06	0.003	0.22	-0.07	-0.38**	-0.31*	0.11	0.1	-0.09	-0.09	-0.39**	-0.12	1	
SN%	0.03	0.15	0.06	-0.31*	-0.11	-0.08	-0.05	-0.09	-0.06	-0.05	-0.14	-0.07	-0.04	0.34**	1

DFL, Days to 50% flowering; CC, SPAD at anthesis stage; LA, Leaf area; PHT, Plant height; PN, Panicle number; GY, Grain yield (g); PW, Panicle weight (g); DSY, Dry stalk yield (g); FSY, Fresh stalk yield (g); LN, Leaf number; NT, Number of tillers; HI, Harvest index; TW, Test weight (100 seed weight); GN%, N content in grain, SN%, N content in stalk. * denotes $p = 0.05$, ** denotes $p = 0.001$.

Population Structure and Diversity Analysis

The hierarchical population structure analysis with a range of $k = 1-8$ sub-populations using 89,770 SNPs, helped identify $k = 3$ (Figure 2A) and accessions contributing to each subpopulation (Figure 2B). The percentage of heterozygous alleles is three. Phylogenetic analysis using the GBS data³ formed three distinctive clusters (grain N and stalk N data), to develop independent phylogenetic trees. The individuals with high and low N values (for both grain and stalk) were randomly

distributed. The largest unit clustered together 13 accessions (in red) for stalk N content ranging from 0.41 to 0.50% followed by a set of 12 accessions (in blue) for the range 0.51–0.60%. For grain N content, the largest cluster consisted of 8 genotypes (in dark blue) with a range of 1.31–1.40%. The other groups 3, 4, and 5 had six genotypes each (in red, orange, and green) with a extensive range of 1.41–1.60% (Figures 3A,B).

Identification of Contrasting Sorghum Genotypes for NUE

A set of 10 accessions each, better and poor performers, was selected based on grain yield data (over two seasons)

³<https://doi.org/10.21421/D2/BYZY2F>

TABLE 5 | Correlation coefficient analysis of 60 genotypes for different traits at N100 condition across seasons.

Trait	DFL	CC	LA	PHT	PN	GY	PW	DSY	FSY	LN	NT	HI	TW	GN%	SN%
DFL	1														
CC	-0.15	1													
LA	0.54**	-0.14	1												
PHT	0.17	-0.21	0.08	1											
PN	-0.28*	-0.12	-0.29*	0.15	1										
GY	-0.05	-0.16	-0.24	-0.02	0.1	1									
PW	0.02	-0.1	-0.26*	-0.19	-0.01	0.84**	1								
DSY	0.51**	-0.29*	0.48**	0.38**	-0.06	-0.19	-0.17	1							
FSY	0.5**	-0.22	0.45**	0.36**	-0.01	-0.06	0.002	0.9**	1						
LN	0.06	-0.22	0.39**	0.13	0.37**	-0.31*	-0.26*	0.16	0.21	1					
NT	-0.34**	-0.17	-0.13	-0.02	0.73**	-0.15	-0.15	-0.08	-0.11	0.49**	1				
HI	-0.3*	-0.01	-0.35**	-0.12	0.02	0.79**	0.66**	-0.64**	-0.5**	-0.35**	-0.15	1			
TW	-0.11	0.05	-0.09	-0.08	-0.42**	0.09	0.15	-0.01	0.003	-0.25	-0.39**	0.02	1		
GN%	0.07	0.22	0.28*	-0.01	-0.1	-0.4**	-0.4**	0.004	-0.005	0.05	0.01	-0.32*	-0.11	1	
SN%	0.07	0.23	-0.01	-0.11	-0.08	-0.21	-0.12	-0.05	0.01	-0.13	-0.18	-0.09	-0.04	0.19	1

DFL, Days to 50% flowering; CC, SPAD at anthesis stage; LA, Leaf area; PHT, Plant height; PN, Panicle number; GY, Grain yield (g); PW, Panicle weight (g); DSY, Dry stalk yield (g); FSY, Fresh stalk yield (g); LN, Leaf number; NT, Number of tillers; HI, Harvest index; TW, Test weight (100 seed weight); GN%, N content in grain, SN%, N content in stalk. * denotes $p = 0.05$, ** denotes $p = 0.001$.

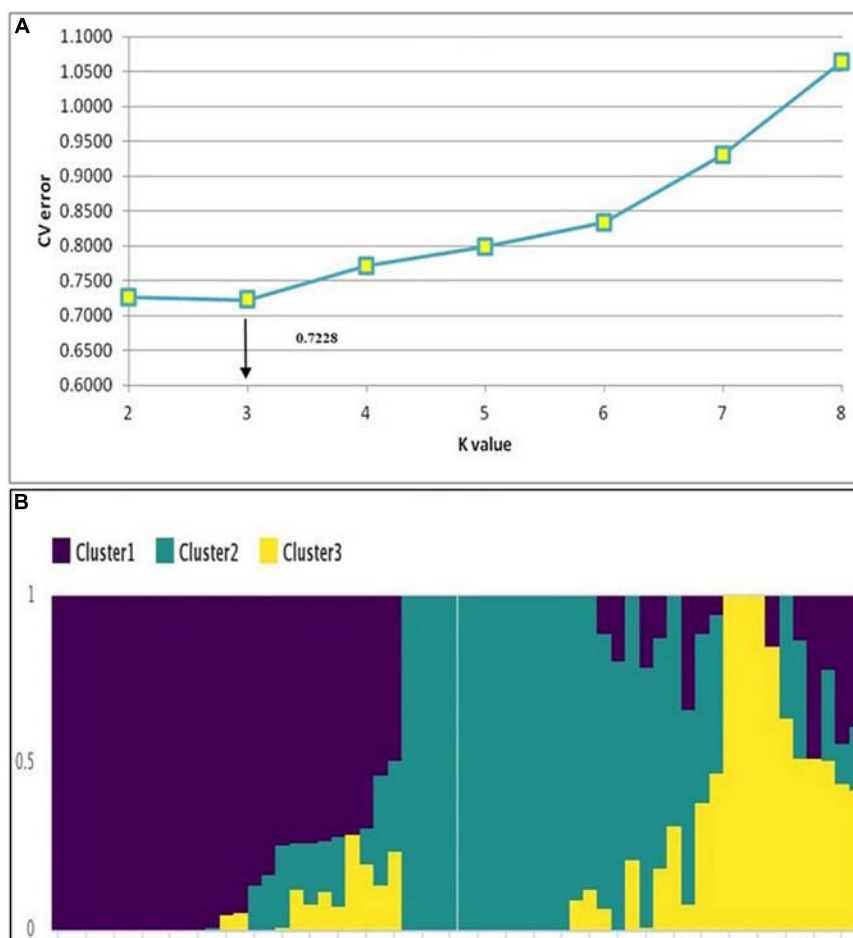
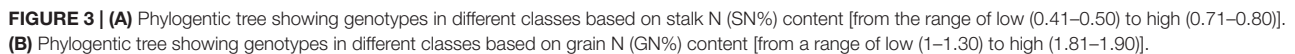


FIGURE 2 | Population structure analysis. **(A)** Rate of change in CV error between successive K-values; K-values ranged from 1 to 8. **(B)** Population structure of the 58 sorghum accessions showing three major clusters.



conditions. These genotypes were identified based on their yield performance in terms of grain, fodder, and both grain and fodder (**Figure 4**).

Expression analysis of key genes associated with N uptake, assimilation, and remobilization under different N conditions,

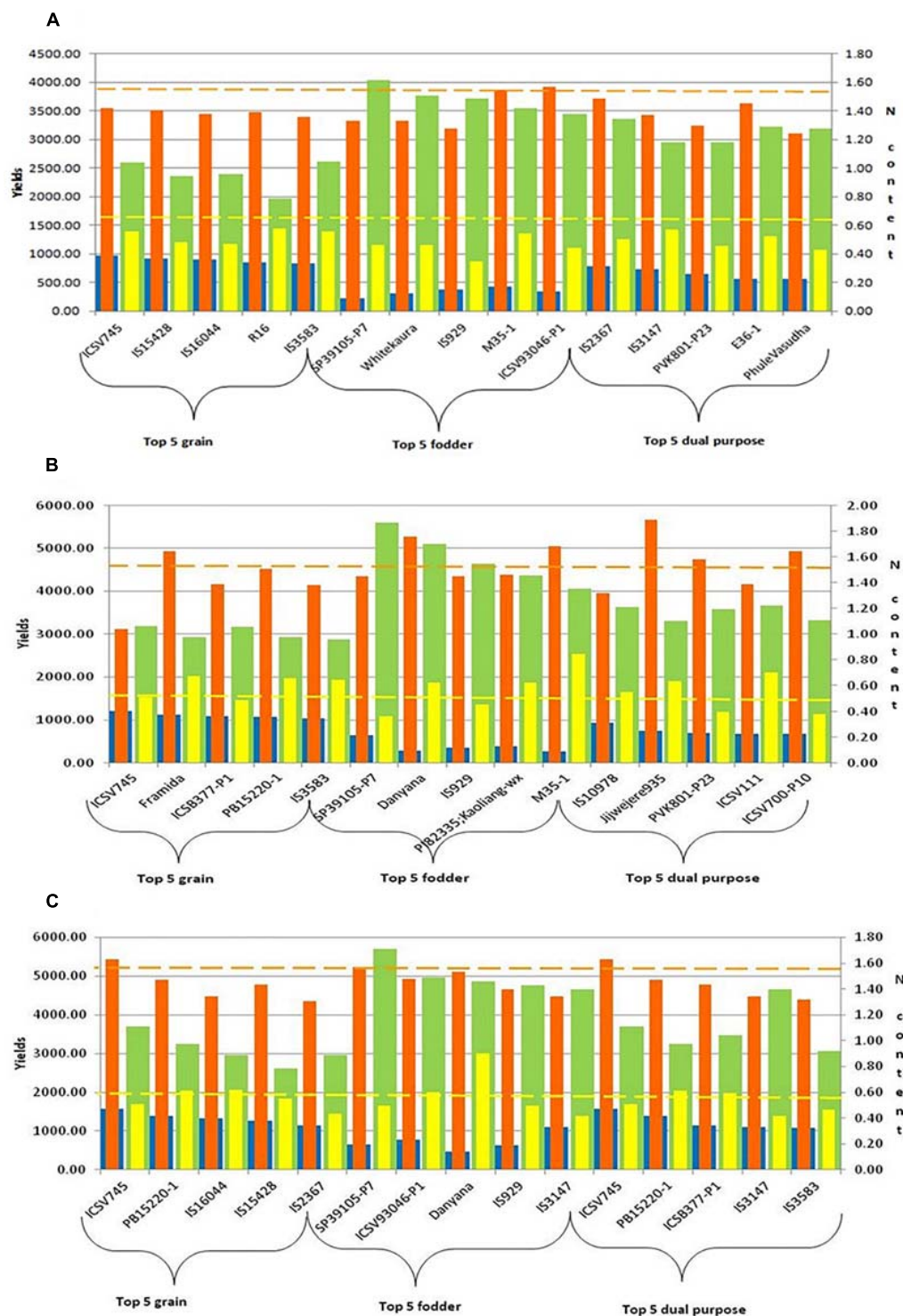


FIGURE 4 | Top five grain, fodder and dual purpose sorghum genotypes under different N (A: N0, B: N50, and C: N100) regimes [Note: ■ GY (g): Grain yield (g); ■ DSY (g): Dry stak yield (g); ■ GN%: N content in grain; ■ SN%: N content in stak].

helps to understand the contrasting response of genes in different tissue samples of high (PVK801) and low (HDW703) NUE genotypes.

Ammonium Transporters

The expression profiles of the *SbAMT* genes (Supplementary Table 3) in shoot tissues, under N0 compared with N100

condition, were down-regulated for both test accessions. In shoot samples of PVK801, a down-regulation of 30, 31, 2, and 65-folds, whereas in HDW703 11, 7, 14, and threefold down-regulation was recorded for *SbAMT1-1*, *SbAMT1-2*, *SbAMT2-1*, and *SbAMT2-2* genes, respectively.

However, the root tissues recorded contrasting regulation than shoot tissues. Root samples of PVK801 exhibited 3, 2, 5, and 4-fold down-regulation of *SbAMT1-1*, *SbAMT1-2*, *SbAMT2-1*, and *SbAMT2-2* genes, respectively; whereas low NUE genotype HDW703 observed up-regulation of *SbAMT1-1*, *SbAMT1-2*, *SbAMT2-1*, and *SbAMT2-2* genes by 11, 15, 14, and 40-folds, respectively (Figure 5).

Nitrate Transporters

Similar to ammonium transporter genes shoot samples of high NUE genotype PVK801 showed 12, 66, 32, and 13-folds down-regulation. In contrast, low NUE genotype HDW703 exhibited

about 2, 10, 4, and 10-folds down-regulation for *SbNRT1-1A*, *SbNRT1-1B*, *SbNRT1-2*, and *SbNRT2-1* genes, respectively in N0 compared to N100 condition.

SbNRT genes exhibited contrasting expression profiles in root samples of high and low NUE genotypes. In high NUE genotype PVK801 about 2, 21, 13, and 3-fold down-regulation was recorded, for *SbNRT1-1A*, *SbNRT1-1B*, *SbNRT1-2*, and *SbNRT2-1* genes, respectively. On the contrary, in HDW703 about 30, 2, 10, and 4-fold up-regulation of *SbNRT1-1A*, *SbNRT1-1B*, *SbNRT1-2*, and *SbNRT2-1* genes was observed (Figure 6).

N Assimilation and Remobilization Related Genes

In shoot samples of high NUE genotype PVK801, N assimilatory and remobilization related genes *SbNiR[NADH]3*, *SbGS*, *SbGOGAT[NADH]*, *SbGOGAT[Fd]*, and *SbGDH* down-regulated by 30, 7, 59, 38, and 3-folds, respectively. For low NUE genotype HDW703 exhibited 11, 7, 14, and 3-fold down-regulation of

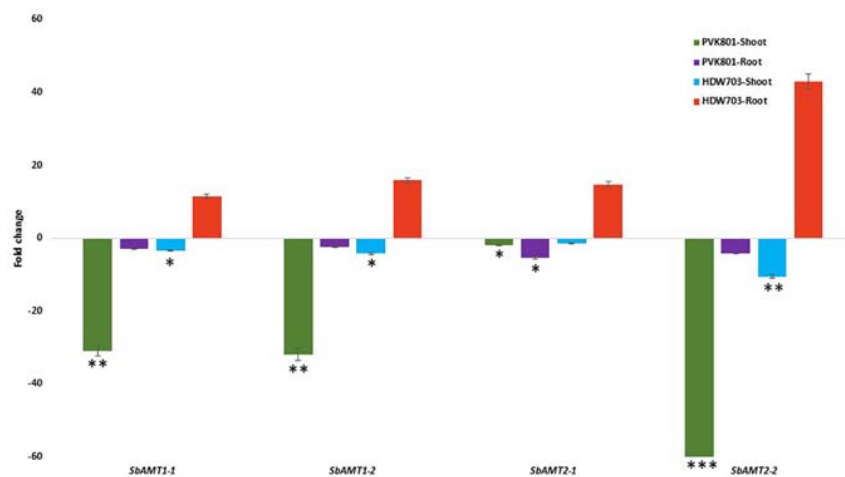


FIGURE 5 | Expression profiles of *SbAMT* genes in shoot and root samples of high (PVK801) and low NUE (HDW703) genotypes, under N0 condition. Here N0 condition taken as treated and N100 used as control. *, **, and *** denotes significance at 5, 1, and 0.1% P, remaining not significant.

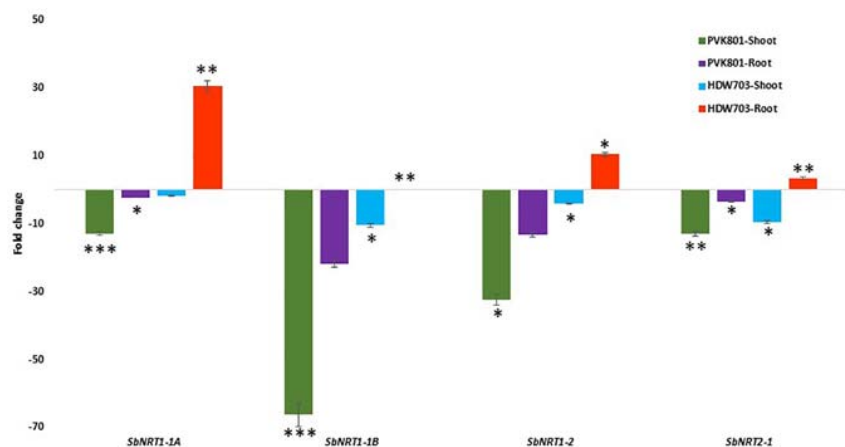


FIGURE 6 | Expression profiles of *SbNRT* genes in shoot and root samples of high (PVK801) and low NUE (HDW703) genotypes, under N0 condition. Here N0 condition taken as treated and N100 used as control. *, **, and *** denotes significance at 5, 1, and 0.1% P, remaining not significant.

SbNiR[NADH]3, *SbGS*, *SbGOGAT[NADH]*, and *SbGOGAT[Fd]* genes respectively. On the contrary to other genes, *SbGDH* exhibited about twofold up-regulation in shoot samples of HDW703 in N0 compared to the N100 condition.

Root samples of PVK801 exhibited down-regulation of all the above genes in the N0 condition. About 23, 3, 7, 4, and 4-fold down-regulation of *SbNiR[NADH]3*, *SbGS*, *SbGOGAT[NADH]*, *SbGOGAT[Fd]*, and *SbGDH* genes in N0 compared to N100 condition was observed. However, contrasting expression of N assimilatory and remobilization related genes in low NUE genotype (HDW703), with about 2, 21, 9, 14, and 63-fold up-regulation of *SbNiR[NADH]3*, *SbGS*, *SbGOGAT[NADH]*, *SbGOGAT[Fd]*, and *SbGDH* genes, respectively, were noted (Figure 7).

DISCUSSION

We conducted a series of field and lab experiments to understand the genotype response, nitrogen dosage effects, and season-specific variations associated with NUE in a set of sixty diverse sorghum accessions. Different physiological, agronomical, and biological yield attributes associated with NUE were systematically studied in three different N dosages (0, 50, and 100% of the recommended N) over two seasons.

With the increase in N dose, the higher SPAD readings were observed. This variation could be of increased availability of N in leaves under N50 and N100 conditions, similar variations were reported in three sorghum varieties of Nigeria (Ajeigbe et al., 2018) and rice (Hassan et al., 2009), wheat (Noulas et al., 2018). This response was also similar to dicotyledons plants (Poorter and Evans, 1998), spinach (Gülser, 2005). An increase in plant height with an increased nitrogen rate application, which might be attributed to effect of N application in plant growth and development. Our results are in accordance with studies in sorghum varieties of east Africa and Nigeria (Shamme and

Raghavaiah, 2016; Ajeigbe et al., 2018) and rapeseed (Khan et al., 2017). Generally physiological maturity of sorghum accelerated under low N than high N conditions, interestingly in this study flowering time showed a minimum response to N doses at field in accordance to few sorghum varieties of Nigeria (Ajeigbe et al., 2018), mung bean (Achakzai et al., 2012) and wheat (Guttieri et al., 2017). Tiller number of the present study increased with enhanced N rates. It is evident that, optimum N availability stimulates tiller number, hence increased number of panicles there by yield improvement. Shamme and Raghavaiah. (2016) and Ajeigbe et al. (2018) also reported higher number of tillers and panicles and weight of panicles with increasing N application in sorghum, in *Brachypodium* (Yang and Udvardi, 2018) and in mung bean (Achakzai et al., 2012) and rice (Pan et al., 2016). There have been reports on N promoting the spikelet number per panicle and thus yield (Yoseftabar, 2013). The analysis of variance revealed significant influence of genotypes, N dosage treatments and their interaction on grain yield which is in agreement with the earlier reports (Gelli et al., 2016; Shamme and Raghavaiah, 2016). Grain yield of sorghum is the final outcome of yield components, more than 40% reduction in grain yield was observed in N0 compared to N100 for individual and across-season of field evaluations. Significant increase in grain yield with the increased N fertilizer application was well established, as N is the main macro nutrient critical for increased yield and related components such as panicle numbers and tiller numbers in major cereals (Gallais and Hirel, 2004; Laperche et al., 2006; Gelli et al., 2016; Shamme and Raghavaiah, 2016; Srikanth et al., 2016). Less than 8% decrease in test weight was observed amongst genotypes grown in N0 compared to N100 conditions in both the seasons, which could be because of increased spikelet number with increase in N application may affect the spikelet size (Qiao et al., 2013). Contrastingly, it was also reported that, test weight is an important trait and yield determining component, reported to be a genetic and also a stable varietal

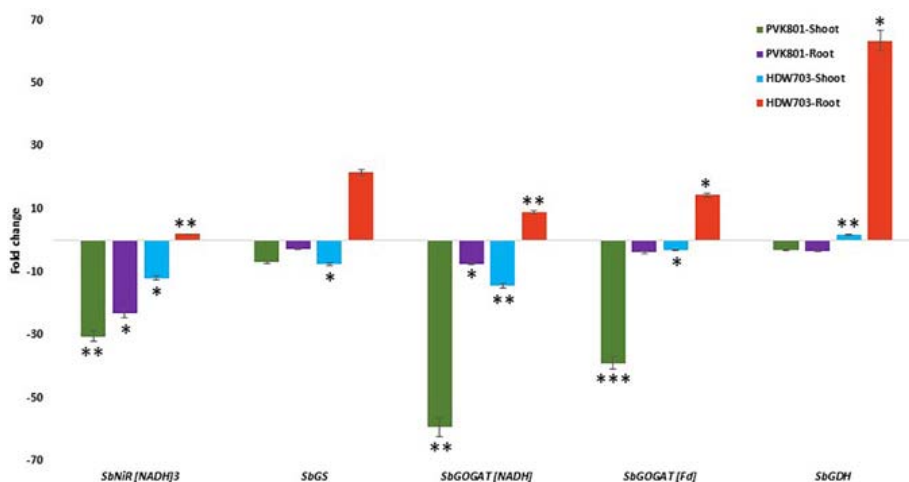


FIGURE 7 | Expression profiles of assimilatory and remobilization related genes in shoot and root samples of high (PVK801) and low NUE (HDW703) genotypes, under N0 condition. Here N0 condition taken as treated and N100 used as control. *, **, and *** denotes significance at 5, 1, and 0.1% P, remaining not significant.

trait mostly depends on spikelet size hence not influenced by environmental factors (Ashraf et al., 1999; Kanfany et al., 2014).

Harvest index is the measure of success in translocation of absorbed assimilates (from source) into economic yields (to sink). Genotypes with more harvest index distribute carbohydrate in to the product efficiently (Li et al., 2012). In the present study increase in the N dosages consistently improved the harvest index, it was significant within genotypes, season and interaction effects of genotypes with seasons indicates the existence of substantial genetic variability for this trait and could be used as selection criterion for improvement of NUE. About 26% reduction in harvest index was observed in N0 conditions in comparison with recommended N (N100). Our results are in line with the earlier studies explaining harvest index increased with the increase in fertilizer N supply in sorghum and other crops (Lawrence et al., 2008; Shamme and Raghavaiah, 2016; He et al., 2017).

Significant variations for grain N content was observed within genotypes and the seasons. Grain N content was gradually increased by increasing fertilizer N applications. About 9 and 17% reduction in grain N content was observed in the present study suggesting the complexity of this trait in different treatments. N content in stover exhibited significant variations in genotypes, treatments, seasons, and their interactions. N content in grain and stover and its association with grain yield and biomass appears to be highly variable across genotypes, seasons, locations, crop duration of genotypes and timing of N application (Zhang et al., 2009; Artacho et al., 2011; Gueye and Becker, 2011; Zhao et al., 2012; Sui et al., 2013; Li et al., 2014; Gelli et al., 2016; Shamme and Raghavaiah, 2016; He et al., 2017).

Correlation Coefficient Analysis

In all the three N treatments, correlation coefficient analysis revealed significant correlations for grain yield, which is positively correlated with panicle weight and harvest index, except across seasons. In a recent study on rapeseed similar correlation trends were observed where genotypes differed significantly in agronomic traits with no consistency in correlation among morphological and N utilization efficiency (He et al., 2017). Interestingly, significant negative correlations were observed for fresh stover weight, dry stover weight, leaf area with N content in grain. Contrary to our results, there are reports in sorghum and other crops describing, grain yield strongly correlated with biomass yield suggesting the role of economic sink strength (Ntanos and Koutroubas, 2002; Gelli et al., 2016; Srikanth et al., 2016). This could be attributed to the differential response of photo-sensitivity and dual-purpose nature of most of the accessions involved in this study. Correlation of the yield and other associated NUE traits appear to be variable with N fertilizer dosage and the genotypes and/or varieties (Sui et al., 2013; Zhang et al., 2013). In the present study, N content in grain positively correlated with panicle weight, grain yield and harvest index under N0 conditions, suggests these traits importance in screening high NUE lines of sorghum. Fresh and dry stover yield positively

correlated with leaf area, plant height, days to 50% flowering in all the treatments and all the conditions. Leaf area correlated with leaf number, dry stover yield in all the treatments of the field trials. N uptake, assimilation and its distribution in vegetative, and reproductive parts are critical processes which determine the grain yield (Guindo et al., 1992). Panicle number significantly correlated with leaf number and number of tillers in all the treatments of the present field study proposes the differential responses of yield and associated traits and their role in plant growth and development under differential nitrogen conditions in sorghum.

Population Structure and Diversity Studies

The sorghum genotypes selected for the study represent the genetic diversity of different countries viz., India, Niger, Sudan, South Africa, Pakistan, Yemen, Cameroon, and the United States depicting the range of genotypes and their origin. The phylogenetic tree from the GBS data formed three distinctive clusters. The high diversity of plant populations may be due to the evolutionary history, genetic drift, and geographic range of the species and their characteristics (Al Salameen et al., 2020). The genotypes representing the grain N and stalk N were randomly distributed over the phylogenetic tree. A similar trend of high genetic variation was observed in sorghum landraces and 44 genotypes selected at random from sorghum mini-core collection (Dossou-Aminon et al., 2015; Satish et al., 2016). Similar overlapping of genotypes in phylogenetic tree was observed in *RhanteriummeppaposumOliv.* (Arfaj). The major reason was due to low genetic distance and differentiation among the populations (Hogbin and Peakall, 1999).

Identification of Contrasting Sorghum Genotypes Under Varying N Conditions

Sorghum is grown across the world for food-fodder-feed-fuel purpose. This is also include industrial use of silage and biofuel applications, especially in developed world. These different end uses are delivered through breeding products developed for specific purpose and with specific cultivation practices in a given agro-ecology. This also include differential N application and N use scenarios. Accelerating the sorghum productivity in terms of grain and fodder yield under intensive agriculture with high dose of N fertilizer application is a cost effective approach for industrial needs. Similarly improving N use efficiency under substance farming (with no or minimal fertilizer inputs) for poor and marginal farmers is possible and also help realizing climate resilience. In this study, apart from product specific (grain, fodder and dual purpose) sorghum accessions (**Figure 4**) in varying N conditions, a set of better and poor performing accessions specifically under low N conditions were identified to investigate the molecular basis for low N tolerance in sorghum. Identification of contrasting (high and low NUE) accessions is the first step toward deciphering the candidate genes and pathways associated with N metabolism, addressing to the different end-uses of sorghum cultivars.

Differential Expression of NUE Associated Genes in Contrasting Sorghum Genotypes Under Varying N Regimes

Using quantitative real-time PCR analysis, expression profiles of N uptake, assimilation and remobilization related genes were studied in shoot and root samples of high and low NUE sorghum genotypes. In these assays N0 condition taken as treated and N100 (recommended N) condition considered as control. Ammonium and nitrate are the major sources of fertilizer N in agricultural soils and act as signal and nutrients for plant growth and development. N fertilizer source nitrate is usually absorbed by plant roots by *NRT1* and *NRT2* (low and high-affinity nitrate transporters), which is further reduced to nitrite by nitrate reductase gene (*NR*), and converted to ammonium by nitrite reductase (*NiR*). Ammonium further assimilated into amino acids by glutamine synthetase (*GS*) and glutamate synthase (*GOGAT*) genes (Crawford, 1995; Lam et al., 1996; Stitt, 1999).

Expression analysis of NUE associated genes under varying N doses in different crops has exhibited differential expression profiles in different tissue samples. Ammonium is one of the readily available form of N for the plants and NH_4^+ uptake in plant roots and its transport to shoots is mediated by ammonium transporters belonging to the *AMT* family. In crop plants, *AMT1*, *NRT2* along with the support of *NRT3* genes act as high-affinity transporters (HATs) in low N situations (Masclaux-Daubresse et al., 2010; Zhou et al., 2016). On the other hand, *AMT2* and *NRT1* works as low affinity transporters (LATs) performs competently under high dose of ammonium and nitrate (Neuhäuser et al., 2007; Wang et al., 2012). HATs and LATs exhibit differential affinities toward N source viz., ammonium and nitrate, it is expected that crops can use a wide range of soil N source applications. In the present study, root samples of low NUE genotype HDW703 exhibited up-regulation of all the *SbAMT* genes (*SbAMT1-1*, *SbAMT1-2*, *SbAMT2-1*, and *SbAMT2-2*) in N0 compared to N100 situation. On the contrary root samples of high NUE genotype PVK801 exhibited down-regulation of all the *SbAMT* genes. In rice (Sonoda et al., 2003), also reported similar results, a constitutive expression profiles of *OsAMT1-1* gene well known to be an important member of HAT sub-family, was reported in shoot and root samples. It was well documented that, *AMT1-1* is one of the key gene and potential candidate for improving NUE, plant growth, and grain yield for both low and optimal fertilizer N situations (Ranathunge et al., 2014).

In shoot samples of PVK801 and HDW703 down-regulation of all the *SbNRT* genes (*SbNRT1-1A*, *SbNRT1-1B*, *SbNRT1-2*, and *SbNRT2-1*) was observed in N0 condition compared to N100. In root samples of high NUE genotype PVK801 all the *SbNRT* genes of the study showed down-regulation in N0 conditions whereas low NUE genotype HDW703 exhibited up-regulation of all the *SbNRT* genes under N0 conditions. Contrasting expression profiles of these genes in high and low NUE genotypes indicate different genetic background effects (including breeding and selection history). This clearly reflects in *NRT* genes responses toward nitrate dosages. Variations in the expression of the *SbNRT* genes in sorghum leaves and roots between N0 and N100 conditions might be varied with the quantitative fold

changes in gene copies also. Enhanced expression of *OsAMT1-1* and *OsNRT2-1* gene was reported in rice seedlings grown in hydroponics conditions using growth media supplemented with low N (Shi et al., 2010). In the roots of sorghum and maize seedlings, *SbNRT1-1A* and *SbNRT1-1B* genes exhibited enhanced expression under N limited conditions (Matsumura et al., 1997; Awada, 2017). Our results are in agreement with the previous study by Fan et al. (2007), Nitrate reductase activity reduced in N sensitive genotype but not altered in NUE rice cultivar under limited fertilizer N condition in hydroponics. One probable reason for the enhanced expression of *NRT* genes in low NUE genotype HDW703 could be increased nitrate accumulation in roots. Another probability could be these *NRT* genes capability to act as potential N sensors that facilitate plants to sense and exploit available nitrate source in soil. It is well characterized that variation for improved N absorption (putatively by *NRT* genes) under low N is contributing to variation in NUE (Masclaux-Daubresse et al., 2010; Shi et al., 2010; Awada, 2017).

It was well established that, N source as a signal induces the differential expression of NUE associated genes including *NRT1*, *NRT2*, *NR*, *NiR*, *GS*, and *GOGAT* (Campbell et al., 1986; Crawford, 1995; Lam et al., 1996; Stitt, 1999). Similar trends were found in our studies with shoot samples of high NUE genotype PVK801 exhibiting down-regulation of *SbNiR*[*NADH*]3, *SbGS*, *SbGOGAT*[*NADH*], *SbGOGAT*[*Fd*], and *SbGDH* genes. Root samples of PVK801 exhibited down-regulation of all the above genes in N0 condition. Low NUE genotype HDW703 exhibited down-regulation of N assimilatory and remobilization related genes (*SbNiR*[*NADH*]3, *SbGS*, *SbGOGAT*[*NADH*], and *SbGOGAT*[*Fd*]) in shoot samples except *SbGDH* genes showed up-regulation. Whereas in root samples of HDW703 all the assimilatory and remobilization associated genes (*SbNiR*[*NADH*]3, *SbGS*, *SbGOGAT*[*NADH*], *SbGOGAT*[*Fd*], and *SbGDH*) up-regulated in N0 compared to N100. Differential expression of *NiR* and *GS* between low and high NUE genotypes indicated these enzymes active involvement in imparting efficient uptake and utilization of fertilizer N (Hakeem et al., 2012). Activities of *OsGS*, *OsGOGAT*, and *OsGDH* in a rice hybrid under five N doses described the role of ammonium assimilation enzymes in grain yield improvement and also NUE (Sun et al., 2012). In rice seedlings, Hirose et al. (1997) demonstrated enhanced expression of *NADH-GOGAT* by inducing with NH_4Cl in root tissues. Sonoda et al. (2003) also got similar results, mRNA accumulation of *NADH-GOGAT* remarkably increased of ammonium induction (after 60 min), while the expression pattern of cytosolic *GS1* (Sakamoto et al., 1989) was constitutive throughout the ammonium induction (Tobin and Yamaya, 2001). Another study in barley demonstrated the expression of *GS* and *GDH* during early seed development stage (Hansen et al., 2009). Grabowska et al. (2012) reported, differential expression patterns of *TsGS1-3* and *TsGS2-1* (low), *TsGDH1* (high) for compensating the low expression levels of *GS* genes in seeds. Similar conclusions were also drawn by Tabuchi et al. (2005). Enhanced expression of these genes in root tissues of low NUE genotype HDW703 of our study explains the role of these genes in nitrogen assimilation.

While crop responsiveness to nitrogen availability depends on both crop variety/genotype and its interaction with the level

of N fertilization (Chardon et al., 2010), the possible solutions for- continued improvement in crop productivity under intensive production system and sustainable crop performance under marginal, resource-poor systems will require optimization of genetic, management and policy interventions. It is thus clear that the different sorghum production ecologies, high N-input (intensive) vs., low/no-input, will need differential genetic (and management) solutions for NUE. This will include appropriate gene constellation and deployment of specific genetic donors through targeted breeding efforts. Additional layer of complexity of crop utilization, viz., grain, forage, or dual-purpose, also require specific inputs in terms of crop modulation and breeding for a specific product. We were able to document this variation (Figures 4A–C) by identifying better germplasm sources for grain, forage and dual-purpose products. This information along with putative candidate genes governing different NUE mechanisms will help to advance the NUE research in sorghum, especially for developing better-targeted breeding for improved productivity and livelihoods of the sorghum growing farming communities across the globe.

CONCLUSION

In the current climate change scenario and continuous global perusal of food security, improving NUE of sorghum for both intensive production ecologies and subsistence ecologies is one of the major goals in sorghum improvement programs. In the high N fertilizer application scenario most commonly observed in the intensive agricultural systems, variation in NUE is primarily triggered by differences in N uptake capacity. In contrast, under limited N conditions such as in agro-ecologies of sorghum production areas in Asia and Africa, NUE variation is driven by changes in N remobilization and utilization efficiency. The screening and identification of robust traits in a complex field evaluation study over two seasons is the first of its kind in NUE research and/or nutrition research. In the present study, different leaf traits, growth traits, panicle and biological yield traits along with NUE traits of 60 diverse sorghum genotypes were systematically evaluated at three N dosages at field level over two seasons. Significant variations were observed for different yield and yield-related traits in 0 and 50% N regimes confirming the availability of genotypic variability in sorghum under low N conditions. There is continuous reduction in the composite yield parameters with the reduction in N application. Correlation coefficient analysis revealed the importance of panicle weight and harvest index in different N environments, and these two traits may be ideal traits for identifying N use efficient genotypes. Contrasting sorghum genotypes identified in this study could be crucial to spot genes for NUE and associated traits. Expression profiling of NUE related genes in shoot and root tissues of contrasting lines (PVK801 and HDW703) raised in varying N conditions (N0 vs., N100) revealed some interesting outcomes. Root tissues of contrasting lines exhibited differential expression profiles for uptake (*SbAMT* and *SbNRT*), assimilation (*SbGS*, *SbGOGAT*[*NADH*], *SbGOGAT*[*Fd*], and *SbNiR*[*NADH*]₃), and amino acid biosynthesis associated genes (*SbGDH*). This study identified some of the promising sorghum genotypes

with buffer capacity for yield under low N conditions. This study also identified better performing grain, stover and dual-purpose sorghum genotypes. These genotypes could be potential candidates for in-depth molecular analysis and deployment into the crop improvement programs to develop high NUE sorghum cultivars.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

RG led the project and developed the framework. RG and SD conceptualized and designed this research strategy. SD contributed to the experimental design, genetic material selection, and genomics studies. GC designed the N-dose treatments and N-response observations. SB, KR, and LR performed field and lab experiments, phenotyping, analyzed, and interpreted the data. AR, AV, RD, and PG carried out statistical and bioinformatics analysis. SB, LR, SD, and RG wrote the manuscript. All authors contributed to editing the MS for publication.

FUNDING

Funding support was provided by the Department of Biotechnology (DBT), Government of India, the Biotechnology and Biological Sciences Research Council (BBSRC), United Kingdom, for the CINTRIN project (BT/IN/UK-VNC/42/RG/2015-16 and BB/N013441/1) to RG, SD, KR and LR, and by the Department of Science and Technology-National-Science and Engineering Research Board (DST-SERB) (PDF/2016/003203) for the National Post-Doctoral Fellowship (NPDF) fellowship to SB. RG, SD, and SB are currently supported by the Global Challenge Research Fund (GCRF) of the BBSRC TIGR2ESS programme (BB/P027970/1). This work was published as part of the CGIAR Research Program on Grain Legumes and Dryland Cereals (CRP-GLDC).

ACKNOWLEDGMENTS

We greatly acknowledge the research technicians for their help in phenotyping and ICRISAT Farm and Engineering Services (FES) section for field management.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.643192/full#supplementary-material>

Supplementary Figure 1 | Weather data of field during the years 2016–17 for the complete crop growth cycle.

Supplementary Figure 2 | Weather data of field during the years 2017–18 for the complete crop growth cycle.

Supplementary Table 1 | List of sorghum genotypes used for NUE field phenotyping experiment.

Supplementary Table 2 | Details of soil properties of the NUE field experiment for the year 2016 at ICRISAT, Patancheru.

Supplementary Table 3 | List of gene specific primers used for studying expression profiles of N uptake assimilatory and remobilization related genes in shoot and root tissues of contrasting sorghum genotypes.

Supplementary Table 4 | Correlation coefficient analysis of 60 genotypes at N0 (A), N50 (B), and N100 (C) dosages for 2016–2017 at ICRISAT, Patancheru.

Supplementary Table 5 | Correlation coefficient analysis of 60 genotypes at N0 (A), N50 (B), and N100 (C) dosages for 2017–2018 at ICRISAT, Patancheru.

REFERENCES

- Achakzai, A. K. K., Rajpar, H., Shah, B. H., and Wahid, M. A. (2012). Effect of nitrogen fertilizer on the growth of mungbean [*Vigna radiata* (L.) Wilczek] grown in Quetta. *Pakistan J. Bot.* 44, 981–987.
- Ajeigbe, H. A., Akinseye, F. M., Ayuba, K., and Jonah, J. (2018). Productivity and Water Use Efficiency of Sorghum [*Sorghum bicolor* (L.) Moench] Grown under Different Nitrogen Applications in Sudan Savanna Zone, Nigeria. *Int. J. Agron.* 2018:20767605818. doi: 10.1155/2018/7676058
- Al Salameen, F., Habibi, N., Al Amad, S., Kumar, V., Dashti, J., Talebi, L., et al. (2020). Genetic diversity analysis of *Rhanterium eppaposum* Oliv. by ISSRs reveals a weak population structure. *Curr. Plant Biol.* 21:100138. doi: 10.1016/j.cpb.2020.100138
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109
- Artacho, P., Meza, F., and Alcalde, J. A. (2011). Evaluation of the ORYZA2000 Rice Growth Model under Nitrogen-Limited Conditions in an Irrigated Mediterranean Environment. *Chil. J. Agricult. Res.* 71, 23–33. doi: 10.4067/s0718-58392011000100003
- Ashraf, A., Khaid, A., and Ali, K. (1999). Effects of seeding rate and density on growth and yield of rice in saline soil. *Pakist. J. Biol. Sci.* 2:860–862.
- Awada, F. (2017). *Assesment of Sorghum Response to Nitrogen Availability Fatima Awada To Cite this Version: HAL*. Paris: UNIVERSITE PARIS -SACLAY.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Campbell, C. A., Biederbeck, V. O., Selles, F., Schnitzer, M., and Stewart, J. W. B. (1986). Effect of Manure and P Fertilizer on Properties of a Black Chernozem in Southern Saskatchewan. *Can. J. Soil Sci.* 66, 601–614. doi: 10.4141/cjss86-060
- Chardon, F., Barthélémy, J., Daniel-Vedele, F., and Masclaux-Daubresse, C. (2010). Natural variation of nitrate uptake and nitrogen use efficiency in arabidopsis thaliana cultivated with limiting and ample nitrogen supply. *J. Exp. Bot.* 61, 2293–2302. doi: 10.1093/jxb/erq059
- Crawford, N. M. (1995). Nitrate: nutrient and Signal for Plant Growth. *Plant Cell* 7:859. doi: 10.2307/3870042
- Diatloff, E., Mace, E. S., Jordan, D. R., Filleur, S., Tai, S., Schmidt, S., et al. (2017). The vegetative nitrogen response of sorghum lines containing different alleles for nitrate reductase and glutamate synthase. *Mole. Breed.* 37:138. doi: 10.1007/s11032-017-0738-1
- Dossou-Aminon, I., Loko, L. Y., Adjatin, A., Ewédjé, E. E. B. K., Dansi, A., Rakshit, S., et al. (2015). Genetic divergence in northern benin sorghum (*Sorghum bicolor* L. Moench) landraces as revealed by agromorphological traits and selection of candidate genotypes. *Scient. World J.* 2015:916476. doi: 10.1155/2015/916476
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379
- Fan, X., Jia, L., Li, Y., Smith, S. J., Miller, A. J., and Shen, Q. (2007). Comparing nitrate storage and remobilization in two rice cultivars that differ in their nitrogen use efficiency. *J. Exp. Bot.* 58, 1729–1740. doi: 10.1093/jxb/erm033
- Frink, C. R., Waggoner, P. E., and Ausubel, J. H. (1999). Nitrogen fertilizer: retrospect and prospect. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1175–1180. doi: 10.1073/pnas.96.4.1175
- Gallais, A., and Hirel, B. (2004). An approach to the genetics of nitrogen use efficiency in maize. *J. Exp. Bot.* 55, 295–306. doi: 10.1093/jxb/erh006
- Gelli, M., Duo, Y., Konda, A. R., Zhang, C., Holding, D., and Dweikat, I. (2014). Identification of differentially expressed genes between sorghum genotypes with contrasting nitrogen stress tolerance by genome-wide transcriptional profiling. *BMC Genomics* 15:179. doi: 10.1186/1471-2164-15-179
- Gelli, M., Konda, A. R., Liu, K., Zhang, C., Clemente, T. E., Holding, D. R., et al. (2017). Validation of QTL mapping and transcriptome profiling for identification of candidate genes associated with nitrogen stress tolerance in sorghum. *BMC Plant Biol.* 17:123. doi: 10.1186/s12870-017-1064-9
- Gelli, M., Mitchell, S. E., Liu, K., Clemente, T. E., Weeks, D. P., Zhang, C., et al. (2016). Mapping QTLs and association of differentially expressed gene transcripts for multiple agronomic traits under different nitrogen levels in sorghum. *BMC Plant Biol.* 16:16. doi: 10.1186/s12870-015-0696-x
- Good, A. G., Shrawat, A. K., and Muench, D. G. (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production. *Trends Plant Sci.* 9, 597–605. doi: 10.1016/j.tplants.2004.10.008
- Grabowska, A., Kwinta, J., and Bielawski, W. (2012). Glutamine synthetase and glutamate dehydrogenase in triticale seeds: molecular cloning and genes expression. *Acta Physiol. Plant.* 34, 2393–2406. doi: 10.1007/s11738-012-1085-9
- Gueye, T., and Becker, H. (2011). Genetic variation in nitrogen efficiency among cultivars of irrigated rice in Senegal. *Sustainable Devel.* 3, 35–43.
- Guindo, D., Wells, B. R., Wilson, C. E., and Norman, R. J. (1992). Seasonal Accumulation and Partitioning of Nitrogen-15 in Rice. *Soil Sci. Soc. Am. J.* 56, 1521–1527. doi: 10.2136/sssaj1992.03615995005600050031x
- Gülser, F. (2005). Effects of ammonium sulphate and urea on NO₃⁻ and NO₂⁻ accumulation, nutrient contents and yield criteria in spinach. *Scientia Horticult.* 106, 330–340. doi: 10.1016/j.scienta.2005.05.007
- Guttieri, M. J., Frels, K., Regassa, T., Waters, B. M., and Baenziger, P. S. (2017). Variation for nitrogen use efficiency traits in current and historical great plains hard winter wheat. *Euphytica* 213, 1–18. doi: 10.1007/s10681-017-1869-5
- Hakeem, K. R., Chandna, R., Ahmad, A., and Iqbal, M. (2012). Physiological and Molecular Analysis of Applied Nitrogen in Rice Genotypes. *Rice Sci.* 19, 213–222. doi: 10.1016/S1672-6308(12)60043-2
- Hansen, T. H., Laursen, K. H., Persson, D. P., Pedas, P., Husted, S., and Schjoerring, J. K. (2009). Micro-scaled high-throughput digestion of plant tissue samples for multi-elemental analysis. *Plant Methods* 5, 1–11. doi: 10.1186/1746-4811-5-12
- Hassan, M. S., Khair, A., Haque, M. M., Azad, A. K., and Hamid, A. (2009). Genotypic variation in traditional rice varieties for chlorophyll content, spad value and nitrogen use efficiency. *Bangladesh J. Agricult. Res.* 34, 505–515. doi: 10.3329/bjar.v34i3.3977
- He, H., Yang, R., Li, Y., Ma, A., Cao, L., Wu, X., et al. (2017). Genotypic variation in nitrogen utilization efficiency of oilseed rape (*Brassica napus*) under contrasting N supply in pot and field experiments. *Front. Plant Sci.* 8:1825. doi: 10.3389/fpls.2017.01825
- Hirose, N., Hayakawa, T., and Yamaya, T. (1997). Inducible Accumulation of mRNA for NADH-Dependent Glutamate Synthase in Rice Roots in Response to Ammonium Ions. *Plant Cell Physiol.* 38, 1295–1297. doi: 10.1093/oxfordjournals.pcp.a029120
- Hogbin, P. M., and Peakall, R. (1999). Evaluation of the contribution of genetic research to the management of the endangered plant *Zieria prostrata*. *Conserv. Biol.* 13, 514–522. doi: 10.1046/j.1523-1739.1999.98182.x
- Kanfany, G., El-Namaky, R., Ndiaye, K., Traore, K., and Ortiz, R. (2014). Assessment of rice inbred lines and hybrids under low fertilizer levels in Senegal. *Sustainability* 6, 1153–1162. doi: 10.3390/su6031153
- Khan, S., Anwar, S., Kuai, J., Ullah, S., Fahad, S., and Zhou, G. (2017). Optimization of nitrogen rate and planting density for improving yield, nitrogen use

- efficiency, and lodging resistance in oilseed rape. *Front. Plant Sci.* 8:532. doi: 10.3389/fpls.2017.00532
- Kimber, C. T. (2000). "Origins of domesticated sorghum and its early diffusion to India and China," in *Sorghum: Origin, History, Technology, and Production*, eds C. W. Smith, and R. A. Frederiksen, (New York: John Wiley & Sons), 3–98.
- Lam, H. M., Coschigano, K. T., Oliveira, I. C., Melo-Oliveira, R., and Coruzzi, G. M. (1996). The molecular-genetics of nitrogen assimilation into amino acids in higher plants. *Annu. Rev. Plant Physiol. Plant Mole. Biol.* 47, 569–593. doi: 10.1146/annurev.arplant.47.1.569
- Laperche, A., Brancourt-Hulmel, M., Heumez, E., Gardet, O., and Le Gouis, J. (2006). Estimation of genetic parameters of a DH wheat population grown at different N stress levels characterized by probe genotypes. *Theoret. Appl. Genet.* 112, 797–807. doi: 10.1007/s00122-005-0176-z
- Lawrence, J. R., Ketterings, Q. M., and Cherney, J. H. (2008). Effect of nitrogen application on yield and quality of silage corn after forage legume-grass. *Agronomy J.* 100, 73–79. doi: 10.2134/agronj2007.0071
- Le Gouis, J., Béghin, D., Heumez, E., and Pluchard, P. (2000). Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. *Europ. J. Agron.* 12, 163–173. doi: 10.1016/S1161-0301(00)00045-9
- Li, M., Zhang, H., Yang, X., Ge, M., Ma, Q., Wei, H., et al. (2014). Accumulation and utilization of nitrogen, phosphorus and potassium of irrigated rice cultivars with high productivities and high N use efficiencies. *Field Crops Res.* 161, 55–63. doi: 10.1016/j.fcr.2014.02.007
- Li, X., Yan, W., Agrama, H., Jia, L., Jackson, A., Moldenhauer, K., et al. (2012). Unraveling the complex trait of harvest index with association mapping in rice (*Oryza sativa* L.). *PLoS One* 7:e29350. doi: 10.1371/journal.pone.0029350
- Mace, E. S., Buhariwalla, H. K., and Crouch, J. H. (2003). A high-throughput DNA extraction protocol for tropical molecular breeding programs. *Plant Mole. Biol. Rep.* 21, 459–460. doi: 10.1007/BF02772596
- Masclaux-Daubresse, C., and Chardon, F. (2011). Exploring nitrogen remobilization for seed filling using natural variation in *Arabidopsis thaliana*. *J. Exp. Bot.* 62, 2131–2142. doi: 10.1093/jxb/erq405
- Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., and Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Anna. Bot.* 105, 1141–1157. doi: 10.1093/aob/mcq028
- Massel, K., Campbell, B. C., Mace, E. S., Tai, S., Tao, Y., Worland, B. G., et al. (2016). Whole genome sequencing reveals potential new targets for improving nitrogen Uptake and utilization in *Sorghum bicolor*. *Front. Plant Sci.* 7:1544. doi: 10.3389/fpls.2016.01544
- Matsumura, T., Sakakibara, H., Nakano, R., Kimata, Y., Sugiyama, T., and Hase, T. (1997). A nitrate-inducible ferredoxin in maize roots: genomic organization and differential expression of two nonphotosynthetic ferredoxin isoproteins. *Plant Physiol.* 114, 653–660. doi: 10.1104/pp.114.2.653
- McCullough, D. E., Girardin, P., Mihajlovic, M., Aguilera, A., and Tollenaar, M. (1994). Influence of N supply on development and dry matter accumulation of an old and a new maize hybrid. *Can. J. Plant Sci.* 74, 471–477. doi: 10.4141/cjps94-087
- Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1982). Analysis and Interpretation of Factors Which Contribute to Efficiency of Nitrogen Utilization 1. *Agron. J.* 74, 562–564. doi: 10.2134/agronj1982.00021962007400030037x
- Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., et al. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci. U. S. A.* 110, 453–458. doi: 10.1073/pnas.1215985110
- Neuhäuser, B., Dynowski, M., Mayer, M., and Ludewig, U. (2007). Regulation of NH₄⁺ transport by essential cross talk between AMT monomers through the carboxyl tails. *Plant Physiol.* 143, 1651–1659. doi: 10.1104/pp.106.094243
- Noulas, C., Herrera, J. M., Tziouvakas, M., and Qin, R. (2018). Agronomic Assessment of Nitrogen Use Efficiency in Spring Wheat and Interrelations with Leaf Greenness Under Field Conditions. *Commun. Soil Sci. Plant Anal.* 49, 763–781. doi: 10.1080/00103624.2018.1431267
- Ntanos, D. A., and Koutroubas, S. D. (2002). Dry matter and N accumulation and translocation for Indica and Japonica rice under Mediterranean conditions. *Field Crops Res.* 74, 93–101. doi: 10.1016/S0378-4290(01)00203-9
- Pan, S., Liu, H., Mo, Z., Patterson, B., Duan, M., Tian, H., et al. (2016). Effects of Nitrogen and Shading on Root Morphologies, Nutrient Accumulation, and Photosynthetic Parameters in Different Rice Genotypes. *Scient. Rep.* 6, 1–13. doi: 10.1038/srep32148
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., et al. (2009). The Sorghum bicolor genome and the diversification of grasses. *Nature* 457, 551–556. doi: 10.1038/nature07723
- Paterson, A. H., Bowers, J. E., Feltus, F. A. (2008). "Genomics of sorghum, a semi-arid cereal and emerging model for tropical grass genomics," in *Genomics of Tropical Crop Plants*, eds P. H. Moore and R. Ming (New York: Springer), 469–482. doi: 10.1007/978-0-387-71219-2_20
- Phytozome.JGI.doe.gov. Phytozome (2019). *Phytozome* [online]. Available online at: <https://phytozome.jgi.doe.gov/pz/portal.html> (Accessed on Dec 16, 2019).
- Poorter, H., and Evans, J. R. (1998). Photosynthetic nitrogen-use efficiency of species that differ inherently in specific leaf area. *Oecologia* 116, 26–37. doi: 10.1007/s004420050560
- Qiao, J., Yang, L., Yan, T., Xue, F., and Zhao, D. (2013). Rice dry matter and nitrogen accumulation, soil mineral N around root and N leaching, with increasing application rates of fertilizer. *Europ. J. Agron.* 49, 93–103. doi: 10.1016/j.eja.2013.03.008
- Ranathunge, K., El-Kereamy, A., Gidda, S., Bi, Y. M., and Rothstein, S. J. (2014). AMT1;1 transgenic rice plants with enhanced NH₄⁺ permeability show superior growth and higher yield under optimal and suboptimal NH₄⁺ conditions. *J. Exp. Bot.* 65, 965–979. doi: 10.1093/jxb/ert458
- Rao, I. S., Neeraja, C. N., Srikanth, B., Subrahmanyam, D., Swamy, K. N., Rajesh, K., et al. (2018). Identification of rice landraces with promising yield and the associated genomic regions under low nitrogen. *Scient. Rep.* 8, 1–13. doi: 10.1038/s41598-018-27484-0
- Sahrawat, K. L., Kumar, G. R., and Murthy, K. V. S. (2002). Sulfuric acid–selenium digestion for multi-element analysis in a single plant digest. *Commun. Soil Sci Plan* 33, 3757–3765.
- Sakamoto, A., Ogawa, M., Masumura, T., Shibata, D., Takeba, G., Tanaka, K. et al. (1989). Three cDNA sequences coding for glutamine synthetase polypeptides in *Oryza sativa* L. *Plant Mol. Biol.* 13, 611–614. doi: 10.1007/BF00027323
- SAS Institute Inc [SAS] (2018). *SAS/STAT® 15.1 User's Guide*. Cary, NC: SAS Institute Inc.
- Satish, L., Shilpha, J., Pandian, S., Rency, A. S., Rathinapriya, P., Ceasar, S. A., et al. (2016). Analysis of genetic variation in sorghum (*Sorghum bicolor* (L.) Moench) genotypes with various agronomical traits using SPAR methods. *Gene* 576, 581–585. doi: 10.1016/j.gene.2015.10.056
- Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Prot.* 3, 1101–1108. doi: 10.1038/nprot.2008.73
- Shamme, S. K., and Raghavaiah, C. V. (2016). Sorghum (*Sorghum bicolor* L.) Growth, Productivity, Nitrogen Removal, N- Use Efficiencies and Economics in Relation to Genotypes and Nitrogen Nutrition in Kellem- Wollega Zone of Ethiopia, East Africa. *Adv. Crop Sci. Technol.* 04, 3–10. doi: 10.4172/2329-8863.1000218
- Shi, W. M., Xu, W. F., Li, S. M., Zhao, X. Q., and Dong, G. Q. (2010). Responses of two rice cultivars differing in seedling-stage nitrogen use efficiency to growth under low-nitrogen conditions. *Plant Soil* 326, 291–302. doi: 10.1007/s11104-009-0007-0
- Sonoda, Y., Ikeda, A., Saiki, S., Von Wirén, N., Yamaya, T., and Yamaguchi, J. (2003). Distinct expression and function of three ammonium transporter genes (OsAMT1;1 - 1;3) in rice. *Plant Cell Physiol.* 44, 726–734. doi: 10.1093/pcp/pcg083
- Srikanth, B., Subhakara Rao, I., Surekha, K., Subrahmanyam, D., Voleti, S. R., and Neeraja, C. N. (2016). Enhanced expression of OsSPL14 gene and its association with yield components in rice (*Oryza sativa*) under low nitrogen conditions. *Gene* 576, 441–450. doi: 10.1016/j.gene.2015.10.062
- Stitt, M. (1999). Nitrate regulation of metabolism and growth. *Curr. Opin. Plant Biol.* 2, 178–186. doi: 10.1016/S1369-5266(99)80033-8
- Sui, B., Feng, X., Tian, G., Hu, X., Shen, Q., and Guo, S. (2013). Optimizing nitrogen supply increases rice yield and nitrogen use efficiency by regulating yield formation factors. *Field Crops Res.* 150, 99–107. doi: 10.1016/j.fcr.2013.06.012
- Sun, Y., Ma, J., Sun, Y., Xu, H., Yang, Z., Liu, S., et al. (2012). The effects of different water and nitrogen managements on yield and nitrogen use efficiency in hybrid rice of China. *Field Crops Res.* 127, 85–98. doi: 10.1016/j.fcr.2011.11.015

- Tabuchi, M., Sugiyama, K., Ishiyama, K., Inoue, E., Sato, T., Takahashi, H., et al. (2005). Severe reduction in growth rate and grain filling of rice mutants lacking OsGS1;1, a cytosolic glutamine synthetase1;1. *Plant J.* 42, 641–651. doi: 10.1111/j.1365-3113X.2005.02406.x
- Taylor, J. R. N., Schober, T. J., and Bean, S. R. (2006). Novel food and non-food uses for sorghum and millets. *J. Cereal Sci.* 44, 252–271. doi: 10.1016/j.jcs.2006.06.009
- Tirol-Padre, A., Ladha, J. K., Singh, U., Laureles, E., Punzalan, G., and Akita, S. (1996). Grain yield performance of rice genotypes at suboptimal levels of soil N as affected by N uptake and utilization efficiency. *Field Crops Res.* 46, 127–143. doi: 10.1016/0378-4290(95)00095-x
- Tobin, A. K., and Yamaya, T. (2001). Cellular compartmentation of ammonium assimilation in rice and barley. *J. Exp. Bot.* 52, 591–604. doi: 10.1093/jxb/52.356.591
- Wang, Z., Yu, C., Liu, X., Liu, S., Yin, C., Liu, L., et al. (2012). Identification of Indica rice chromosome segments for the improvement of Japonica inbreds and hybrids. *Theoret. Appl. Genet.* 124, 1351–1364. doi: 10.1007/s00122-012-1792-z
- Yang, J., and Udvardi, M. (2018). Senescence and nitrogen use efficiency in perennial grasses for forage and biofuel production. *J. Exp. Bot.* 69, 855–865. doi: 10.1093/jxb/erx241
- Yoseftabar, S. (2013). Effect Nitrogen Management on Panicle Structure and Yield in Rice (*Oryza sativa* L). *Int. J. Agricult. Crop Sci.* 5, 1224–1227.
- Youngquist, J. B., Bramel-Cox, P., and Maranville, J. W. (1992). Evaluation of Alternative Screening Criteria for Selecting Nitrogen-Use Efficient Genotypes in Sorghum. *Crop Sci.* 32, 1310–1313. doi: 10.2135/cropsci1992.0011183x003200060002x
- Zhang, J. H., Liu, J. L., Zhang, J. B., Cheng, Y. N., and Wang, W. P. (2013). Nitrate-Nitrogen Dynamics and Nitrogen Budgets in Rice-Wheat Rotations in Taihu Lake Region, China. *Pedosphere* 23, 59–69. doi: 10.1016/S1002-0160(12)60080-0
- Zhang, L., Lin, S., Bouman, B. A. M., Xue, C., Wei, F., Tao, H., Yang, X., Wang, H., Zhao, D., and Dittert, K. (2009). Response of aerobic rice growth and grain yield to N fertilizer at two contrasting sites near Beijing, China. *Field Crops Res.* 114, 45–53. doi: 10.1016/j.fcr.2009.07.001
- Zhao, S. -P., Zhao, X., and Shi, W. (2012). Genotype variation in grain yield response to basal N fertilizer supply among different rice cultivars. *Afric. J. Biotechnol.* 11, 12298–12304. doi: 10.5897/ajb12.787
- Zhou, B., Elazab, A., Bort, J., Sanz-Sáez, A., Nieto-Taladriz, M. T., Serret, M. D., et al. (2016). Agronomic and physiological responses of Chinese facultative wheat genotypes to high-yielding Mediterranean conditions. *J. Agricult. Sci.* 154, 870–889. doi: 10.1017/S0021859615000817

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Bollam, Romana, Rayaprolu, Vemula, Das, Rathore, Gandham, Chander, Deshpande and Gupta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Deciphering Genotype-By-Environment Interaction for Target Environmental Delineation and Identification of Stable Resistant Sources Against Foliar Blast Disease of Pearl Millet

OPEN ACCESS

Edited by:

Wolfram Weckwerth,
University of Vienna, Austria

Reviewed by:

Devanna B. N.,
National Rice Research Institute
(ICAR), India
Estefania Carrillo-Perdomo,
INRA Centre Dijon Bourgogne
Franche-Comté, France

*Correspondence:

S. P. Singh
sumerpalsingh@yahoo.com
Rakesh K. Srivastava
r.k.srivastava@cgiaar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 20 January 2021

Accepted: 26 March 2021

Published: 17 May 2021

Citation:

Sankar SM, Singh SP, Prakash G, Satyavathi CT, Soumya SL, Yadav Y, Sharma LD, Rao AR, Singh N and Srivastava RK (2021) Deciphering Genotype-By-Environment Interaction for Target Environmental Delineation and Identification of Stable Resistant Sources Against Foliar Blast Disease of Pearl Millet. *Front. Plant Sci.* 12:656158. doi: 10.3389/fpls.2021.656158

S. Mukesh Sankar¹, S. P. Singh^{1*}, G. Prakash², C. Tara Satyavathi³, S. L. Soumya¹, Yashpal Yadav⁴, L. D. Sharma⁵, A. R. Rao⁶, Nirupma Singh¹ and Rakesh K. Srivastava^{7*}

¹ Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India, ² Division of Plant Pathology, ICAR-Indian Agricultural Research Institute, New Delhi, India, ³ ICAR-All India Coordinated Research Project on Pearl Millet, Jodhpur, India, ⁴ CCS Haryana Agricultural University, College of Agriculture, Bawal, India, ⁵ Rajasthan Agricultural Research Institute, Jaipur, India, ⁶ CABIN, ICAR- Indian Agricultural Statistical Research Institute, New Delhi, India, ⁷ International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India

Once thought to be a minor disease, foliar blast disease of pearl millet, caused by *Magnaporthe grisea*, has recently emerged as an important biotic constraint for pearl millet production in India. The presence of a wider host range as well as high pathogenic heterogeneity complicates host-pathogen dynamics. Furthermore, environmental factors play a significant role in exacerbating the disease severity. An attempt was made to unravel the genotype-by-environment interactions for identification and validation of stable resistant genotypes against foliar blast disease through multi-environment testing. A diversity panel consisting of 250 accessions collected from over 20 different countries was screened under natural epiphytotic conditions in five environments. A total of 43 resistant genotypes were found to have high and stable resistance. Interestingly, most of the resistant lines were late maturing. Combined ANOVA of these 250 genotypes exhibited significant genotype-by-environment interaction and indicated the involvement of crossover interaction with a consistent genotypic response. This justifies the necessity of multi-year and multi-location testing. The first two principal components (PCs) accounted for 44.85 and 29.22% of the total variance in the environment-centered blast scoring results. Heritability-adjusted genotype plus genotype \times environment interaction (HA-GGE) biplot aptly identified “IP 11353” and “IP 22423, IP 7910 and IP 7941” as “ideal” and “desirable” genotypes, respectively, having stable resistance and genetic buffering capacity against this disease. Bootstrapping at a 95% confidence interval validated the recommendations of genotypes. Therefore, these genotypes can be used in future resistance breeding programs in pearl millet. Mega-environment delineation and desirability index suggested Jaipur as the ideal environment for precise

testing of material against the disease and will increase proper resource optimization in future breeding programs. Information obtained in current study will be further used for genome-wide association mapping of foliar blast disease in pearl millet.

Keywords: pearl millet, *Magnaporthe*, blast disease, genotype-environment interaction, heritability, GGE biplots, mega-environments, mixed-model analysis

INTRODUCTION

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is a major climate resilient cereal crop that is cultivated extensively on resource-poor marginal lands of arid and semiarid regions of Asia and sub-Saharan Africa (Anuradha et al., 2017). It forms a source of food and fodder and ensures food and nutritional security to the inhabitants who are practicing low-input agriculture (Pankaj et al., 2020). Being a “Nutri-cereal” and thriving well in any cropping system, it shows a crucial role in defeating malnutrition and improving the socioeconomic status of resource-poor farmers (Govindaraj et al., 2020). Foliar blast (FB) disease of pearl millet caused by the fungus *Pyricularia grisea* (Cooke) Sacc. [Teleomorph: *Magnaporthe grisea* (Herbert) Barr], a disease of negligible importance in past years, has become a severe menace to successful pearl millet cultivation worldwide (Sharma et al., 2018). It is widespread in the different pearl millet-growing ecologies of India but became a very serious threat in both A₁ (includes rainfed areas of western Rajasthan, as well as parts of Gujarat and Haryana, where annual precipitation is anticipated to be < 400 mm and pearl millet productivity is supposed to be less than 100 kg/ha) and A zones (includes North Indian regions excluding regions covered in A₁ zones with an annual rainfall of more than 400 mm), where early- to medium-maturing cultivars are preferred (AICPMIP, 2020). In fact, the disease has reached a critical stage that necessitates a multifaceted approach to its effective management (Sharma et al., 2020).

The disease starts out as a small speck or lesion that grows larger and necrotic, causing widespread chlorosis and premature death of young leaves (**Figure 1**). Lesions can appear as diamond-shaped white to gray or reddish-brown lesions near the leaf tips or margins, or both with reddish to brown borders that extend down and may enlarge, coalesce, and kill entire leaves. During humid weather, particularly with dense plant stands, this disease becomes more severe. *M. grisea* is a seed-borne fungus that often survives in the soil/leaf debris as chlamydospores or free saprophytic mycelium, providing a source of primary inoculums. Repeated infection in a single crop season happens through the dissemination of asexual spore called conidia. FB on pearl millet was observed to be inversely related to days to maturity, green and dry fodder yield, seed yield per plant, and digestive dry matter, influencing crop productivity and quality (Nayaka et al., 2017).

Foliar blast in pearl millet is a multi-cycle disease, and usually crop is grown by resource-poor farmers. Thus, chemical control with repeated sprays in one crop season is not practically feasible. The development of FB-resistant cultivars is a major thrust area for the pearl millet research and development sector

worldwide, as it offers an economic and eco-friendly option for managing the disease. However, due to a limited knowledge of its inheritance (Gupta et al., 2012 and Singh et al., 2018), race specificity (Sharma et al., 2021), and the rapid shift in pathogenicity of the blast fungus, as well as a scarcity of stable resistance in exotic or adapted germplasm (Sharma et al., 2020), progress in transmitting stable resistance to commercial cultivars has been slow. Another most important challenge is poor repeatability in field-plot and greenhouse ratings due to genotype-by-environment interactions (GEIs) (Prakash et al., 2016). The role of genotype \times environment ($G \times E$) interaction is also crucial for the eventual appraisal of sources of durable resistance (Singh et al., 2020). Thus, an identification of resistance source against FB in pearl millet, followed by an appraisal of the durability of resistance and its utilization in resistance breeding programs, is necessary.

Delineation of target testing environment that has good discrimination power, representativeness, and high desirability index is also indispensable for facilitating proper selection of resistance sources as well as curtailing the use of non-informative testing locations, thus reducing the cost of multi-location trials (MLTs). Previous reports have stated that genotype and identification of a testing location would be meaningful within a mega-environment (ME) (Yan et al., 2007). Data collected over years are crucial for defining an ME and improving breeding efficiency (Yan and Holland, 2010). Several statistical tools are available for establishing the role of GEI in the identification of desirable genotypes with specific and broad-spectrum adaptability over different locations (Yan and Kang, 2003). The GGE (genotype plus GEI) biplot methodology usually applies the concept of indirect selection, removes the main effect of the environment, considers only the genotypic main effect with the GEI effect in MLT datasets, and represents the result in graphical mode (Yan et al., 2000). Based on different scaling methods, various forms of GGE biplot have been developed. Among these, the heritability-adjusted GGE (HA-GGE) biplot is the most logical and precise method for the identification of genotypes and test environments (Yan and Holland, 2010).

However, information on the identification of durable FB resistance sources based on multi-environment data is scanty. Therefore, in the present study, we employed the HA-GGE biplot method to evaluate the effects of genotype, environment, and GEI for FB resistance by deploying multi-location and multi-year datasets for detecting ideal genotypes with durable resistance. The recommendation of genotypes for a specific environment was corroborated by bootstrapping at the 95% confidence limit (CL). Additionally, test environments were evaluated in terms of discrimination power, representativeness, and desirability index,

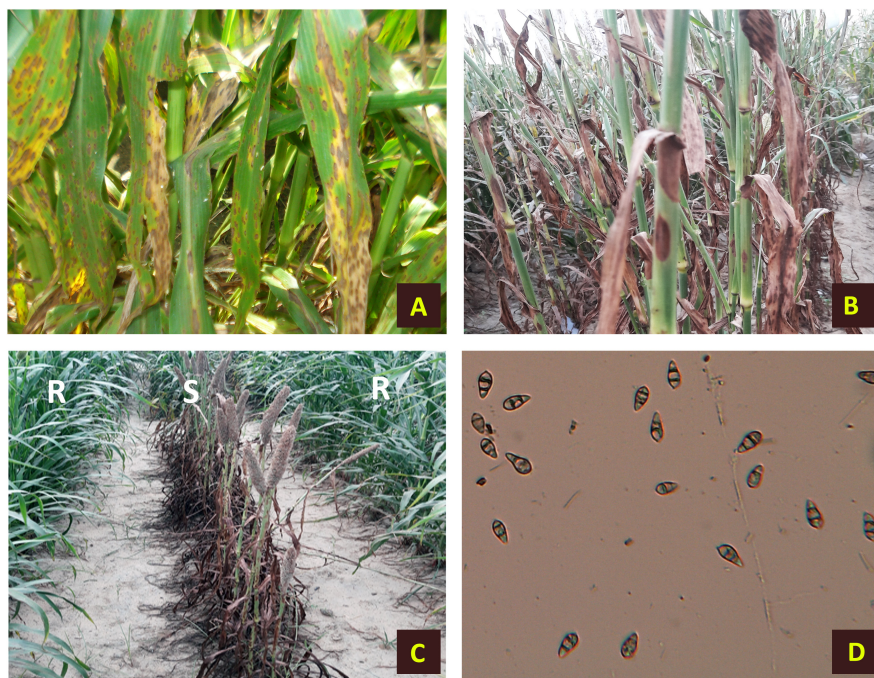


FIGURE 1 | Symptoms of foliar blast on infected pearl millet plants. **(A)** Infected leaves. **(B)** Sheath and stem infection. **(C)** Infected plants along with resistant lines. **(D)** Conidia of *Magnaporthe grisea*.

followed by ME detection, to exclude redundant testing locations and to minimize the cost incurred on future evaluation programs.

MATERIALS AND METHODS

Plant Material and Multi-Environment Field Trials

The experimental materials include a set of diversity panel, which is a subset of PMiGAP (Pearl Millet inbred Germplasm Association Panel). It is denoted as “G” followed by serial number. This panel is composed of 250 accessions collected from over 20 different countries representing the global genetic diversity of pearl millet. The panel includes inbred lines, landraces, released cultivars, germplasm accessions, and advanced breeding lines. These genotypes were evaluated for FB disease under natural epiphytotic conditions in an alpha lattice design with two replications. **Supplementary Table 1** represents the list of genotypes along with details such as subspecies, botanical variety, market type, origin, and pedigree. The panel was evaluated at IARI-New Delhi (28°70'N, 76°58'E, 266.0 MSL) for three seasons (Kharif-2017, Kharif-2018, and Kharif-2019), at CCS-HAU, Bawal (28°07'N, 77°10'E, 288.0 MSL) and at RARI, Jaipur (26°50'N, 75°47'E, 390.0 MSL) for a single season (Kharif-2019). Weather parameters from each environment during tillering to hard dough stage of plant [30–60 days after sowing (DAS)] are presented in **Table 1**. All climatic parameters except rainfall are presented as means over the crop-critical growing period 30–60 DAS. Rainfall is measured as cumulative rainfall

received in mm and the cumulative number of rainy days. A canonical correspondence analysis (CCA) was performed to assess the impact of various environmental factors on FB score. The climatic determinants used for CCA includes the following: maximum and minimum, temperature (Max. Temp and Min. Temp), percentage maximum and minimum, relative humidity (Max. RH and Min. RH), and cumulative rainfall and rainy days (cumulative no. of days when daily rainfall measure above 2.5 mm). Weather data for the present analysis were obtained from the Division of Agricultural Physics, ICAR-IARI, New Delhi, India, and AICRP on Agro-meteorology, Hyderabad. The analysis was carried out in R package-“vegan”.

Agronomic Practices, Disease Screening, and Data Recording in Multi-Environment Field Trials

Each genotype was sown in a plot of two rows each of 2-m length having 65-cm row spacing and a 12-cm plant-to-plant spread. The sites for the research were carefully chosen based on the prevalence of *Magnaporthe grisea*. Spreader rows of FB-susceptible check (ICMB 95444) were planted after every 10th treatment of the test populations, and five rows of the spreader row of the susceptible check were planted on all the sides of the experimental area for maintaining sufficient disease pressure under natural condition. For a normal and healthy crop, standard cultivation practices recommended for pearl millet were regularly followed. FB scores were recorded from five randomly selected representative plants of all the genotypes in each replication, while days to 50% flowering (DFF) were recorded on a plot basis.

TABLE 1 | Geographical identity and climate variables at test environment in respect of temperature, relative humidity (RH), total rainfall, and rainy days along with mean foliar blast (FB) score during the critical period of pearl millet crop growth.

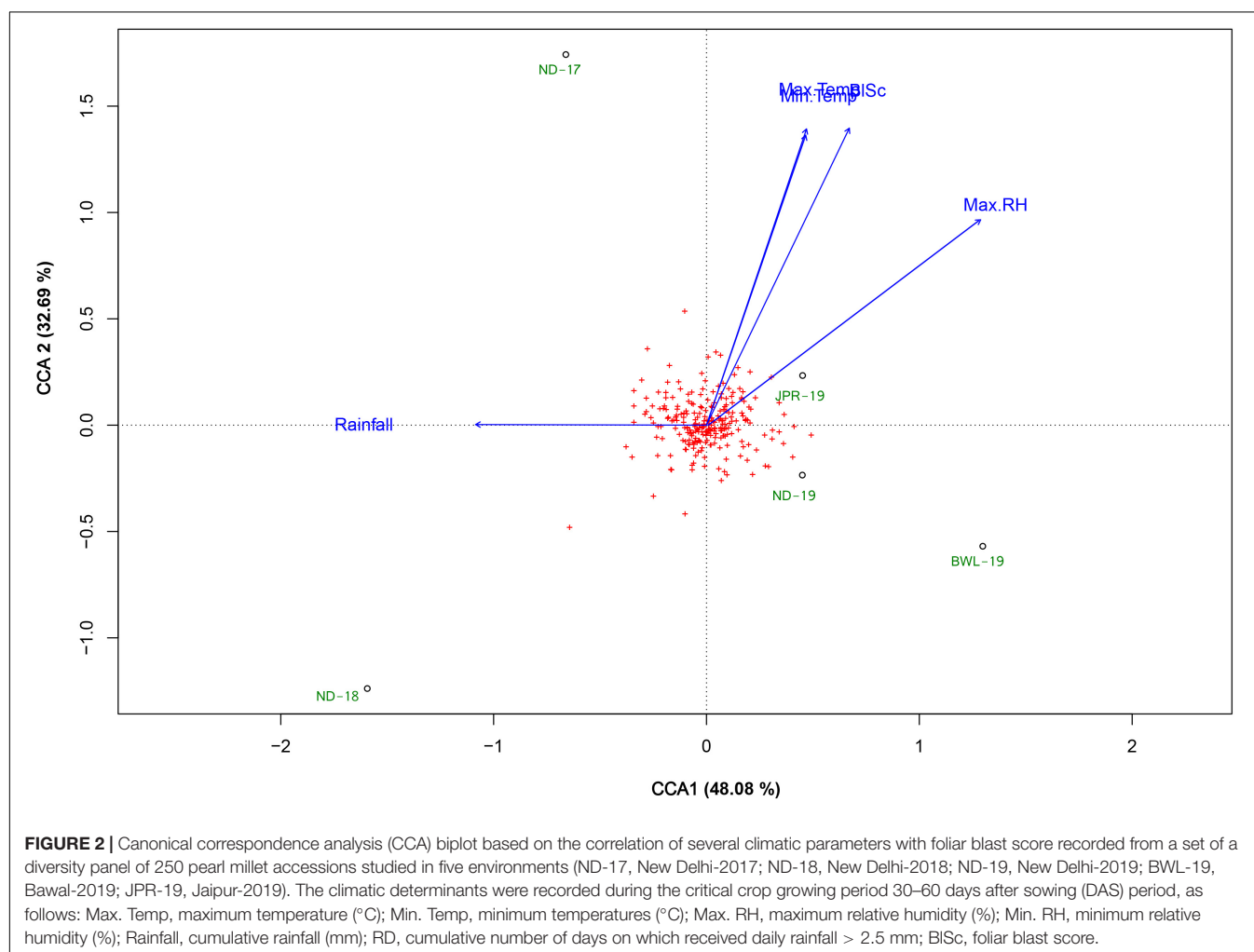
Environment	Latitude	Longitude	Altitude	Max. temp. (°C)	Min. Temp. (°C)	Rainfall (mm)	Max. RH (%)	Min. RH (%)	Rainy days	Mean blast score
New Delhi-17	28.7	76.6	266	33.9	25.2	164.2	92.5	71.5	9.0	4.57
New Delhi-18	28.7	76.6	266	32.0	21.8	188.0	89.4	64.8	11.0	4.30
New Delhi-19	28.7	76.6	266	33.8	25.2	9.0	87.7	62.3	2.0	4.51
Jaipur-19	26.5	75.5	390	32.8	23.2	105.0	82.2	58.4	6.0	4.45
Bawal-19	28.1	77.1	288	33.0	23.6	28.0	90.0	55.4	1.0	4.48

Standard statistical methods were followed for data analysis. The disease was assessed following the 0–9 scale of Prakash et al. (2016) and Nayaka et al. (2017) as described earlier. The GE table of FB mean scores is transformed by subtracting each mean score from 9. Thus, the new score obtained as a consequence of this transformation adopted the same general interpretability principles as yield and other related traits, in which high values are preferred. The genotypes screened were categorized into five groups based on disease scoring: highly resistant (9), resistant (6–8), moderately resistant/susceptible (4–5), susceptible (2–3), and highly susceptible (0–1).

Statistical Analysis

Variance Components and Estimation of Broad-Sense Heritability

Individual and combined analyses of variance (ANOVAs) were conducted on replicated data obtained in different environments (a combination of locations and years). The restricted maximum likelihood (REML) analysis was carried out for each environment, with replications as a fixed effect and genotypes and blocks within replication as random effects, while environments and replications within environments were considered as fixed effects whereas genotypes and genotype



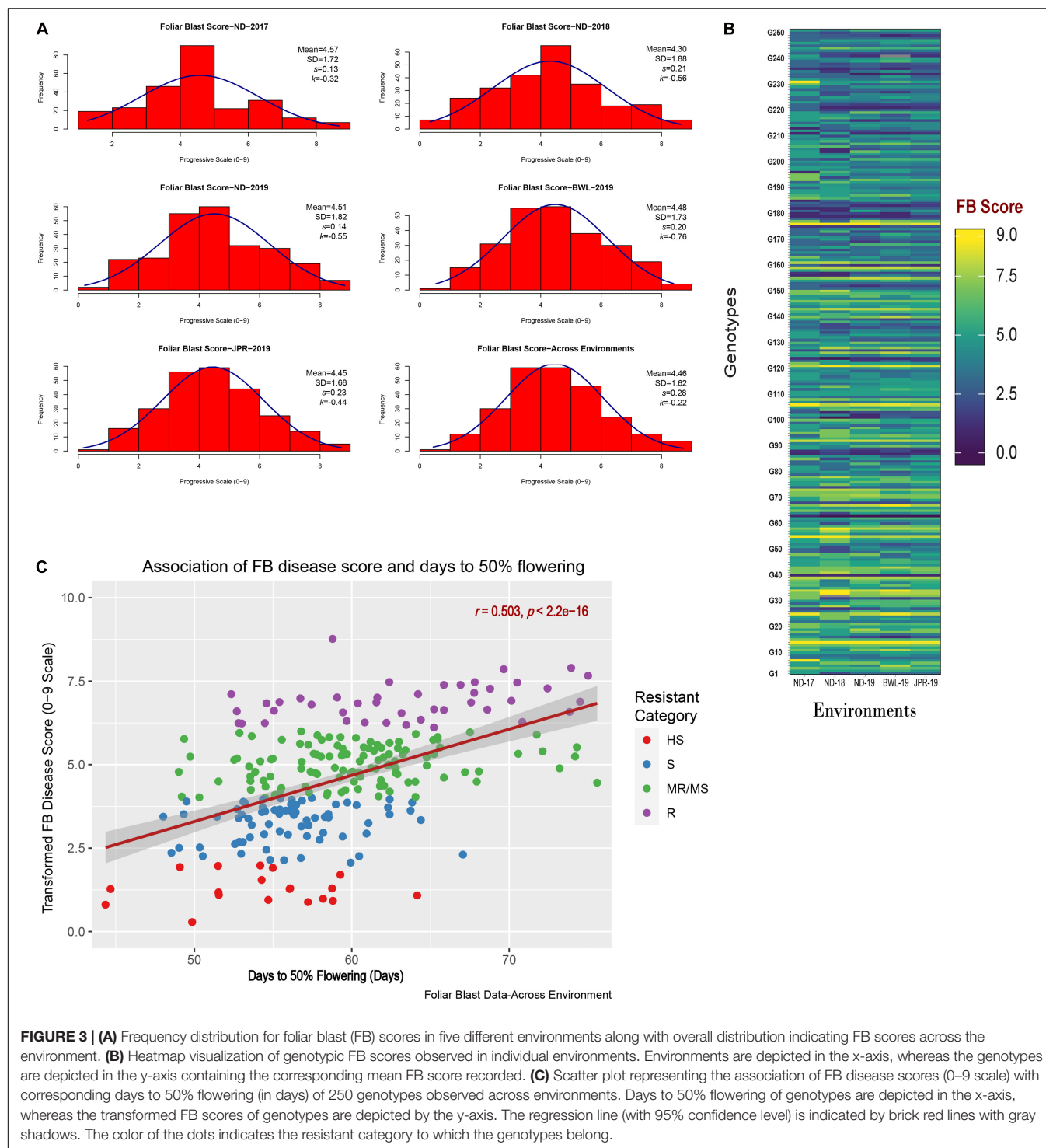


FIGURE 3 | (A) Frequency distribution for foliar blast (FB) scores in five different environments along with overall distribution indicating FB scores across the environment. **(B)** Heatmap visualization of genotypic FB scores observed in individual environments. Environments are depicted in the x-axis, whereas the genotypes are depicted in the y-axis containing the corresponding mean FB score recorded. **(C)** Scatter plot representing the association of FB disease scores (0–9 scale) with corresponding days to 50% flowering (in days) of 250 genotypes observed across environments. Days to 50% flowering of genotypes are depicted in the x-axis, whereas the transformed FB scores of genotypes are depicted by the y-axis. The regression line (with 95% confidence level) is indicated by brick red lines with gray shadows. The color of the dots indicates the resistant category to which the genotypes belong.

interactions with environments were considered as random effects in the REML model for combined environment analysis. The error variances of individual environments (a combination of locations and years) were accounted for combined analysis using the mixed model methodology. Error variance modeling using mixed model analysis takes care of heterogeneous error variances of the individual environment during pooled

analysis. The REML (Patterson and Thompson, 1971) estimation technique was used to estimate three variance components (σ^2_g , σ^2_{ge} , and σ^2_e) for transformed FB score and DFF using the lmer function of the lme4 R-package (Bates et al., 2015). The rand function of the lmerTest package used the likelihood ratio test (LRT) at 5% probability to confirm the significance of the random effects (Kuznetsova et al., 2015).

TABLE 2 | List of genotypes outperforming resistant checks for foliar blast resistance across environments.

Sl.no.	Entry	Code	DFF	FB scores at individual environments					Overall mean	Stability index (ASV)**	Frequency in top 10 resistant line group
				ND-17	ND-18	ND-19	BWL-19	JPR-19			
1	IP 11353	G-62	59	8.2	8.6	8.8	8.5	8.8	8.8	0.05	5/5
2	IP 22423	G-39	74	7.4	8.1	7.8	7.6	7.9	7.8	0.08	5/5
3	IP 7941	G-220	70	7.4	8.1	7.7	7.6	7.7	7.7	0.09	5/5
4	IP 7910	G-123	75	7.3	7.5	7.8	7.6	7.5	7.5	0.07	3/5
5	IP 8280	G-86	62	7.6	6.9	7.8	7.2	7.4	7.4	0.20	3/5
6	IP 12322	G-87	71	6.9	7.2	7.3	7.6	7.8	7.4	0.16	2/5
7	IP 11275	G-178	68	7.5	7.7	7.7	6.7	7.1	7.3	0.24	3/5
8	IP 16082	G-15	66	7.0	7.4	7.5	7.6	6.9	7.3	0.07	1/5
9	IP 12364	G-235	72	7.4	6.8	7.5	7.1	7.1	7.2	0.21	1/5
10	IP 15857	G-233	67	6.5	6.8	8.1	7.1	7.7	7.2	0.27	2/5
11	IP 21169	G-156	69	7.2	8.1	6.8	6.7	6.9	7.1	0.38	1/5
12	IP 5438	G-240	68	6.6	6.8	7.7	NG	7.3	7.1	0.18	2/5
13	IP 8182	G-221	64	6.5	7.2	7.3	7.2	6.8	7.0	0.08	0/5
14	IP 7953	G-159	52	6.2	7.0	7.2	7.1	7.5	7.0	0.22	1/5
15	IP 3122	G-102	59	6.8	7.6	6.7	6.6	6.8	6.9	0.25	1/5
16	IP 9969	G-182	57	7	6.0	7.7	6.7	7.0	6.9	0.33	0/5
17	IP 3106	G-236	62	6.3	7.0	6.5	7.6	7.2	6.9	0.27	1/5
18	IP 9710	G-100	62	6.5	7.6	7.9	4.9	7.0	6.8	0.48	2/5
19	IP 14439	G-174	75	7.2	7.4	6.9	6.9	5.4	6.8	0.28	0/5
20	IP 18293	G-210	70	7.3	4.1	7.8	7.6	7.4	6.8	1.04	3/5
21	IP 10543	G-181	68	7.5	8.1	6.0	5.8	6.3	6.7	0.68	2/5
22	SOSAT-C88	G-179	55	7.4	6.7	7.3	6.7	5.5	6.7	0.30	1/5
23	IP 11584	G-135	60	6.5	7.2	6.3	6.7	6.9	6.7	0.12	0/5
24	IP 6882	G-183	57	6.7	7.2	6.3	6.7	6.6	6.7	0.14	0/5
25	IP 12020	G-241	55	6.1	5.9	7	7.6	7.1	6.7	0.51	1/5
26	IP 3471	G-126	65	6.2	6.3	6.8	6.8	6.8	6.6	0.18	0/5
27	IP 17632	G-26	53	7.0	7.4	7.2	4.5	6.4	6.5	0.69	0/5
28	IP 22419	G-151	66	6.1	7.0	7.2	5.8	6.6	6.5	0.17	0/5
29	IP 10820	G-67	59	6.2	7.2	6.9	5.8	6.2	6.5	0.26	0/5
30	IP 13840	G-226	74	6.6	6.4	6.8	6.5	6.1	6.5	0.13	0/5
31	IP 6098	G-237	69	6.2	7	6.4	6.4	6.5	6.5	0.11	0/5
32	ICMR 11009 (Res. check)	G-248	55	5.7	6.0	6.3	7.5	7.1	6.5	0.49	1/5
33	IP 10761	G-177	63	6.0	6.3	5.8	7.0	7.1	6.4	0.20	0/5
34	IP 11961	G-219	64	5.7	7.2	6.7	5.2	6.5	6.3	0.36	0/5
35	IP 7952	G-130	71	6.0	5.0	7.3	6.1	6.7	6.2	0.48	0/5
36	WC-C75	G-168	61	5.7	5.9	7.0	6.0	6.5	6.2	0.18	0/5
37	IP 7886	G-138	53	4.0	6.2	6.8	7.2	6.8	6.2	0.75	0/5
38	IPC 804	G-122	53	5.6	4.5	7.1	6.9	6.8	6.2	0.73	0/5
39	IP 8074	G-155	60	7.3	6.7	5.4	5.3	6.2	6.2	0.58	0/5
40	ICMR 11003 (Res. check)	G-247	54	5.8	6.0	5.9	6.7	6.5	6.2	0.25	0/5
41	IP 19405	G-204	62	5.9	6.7	6.3	5.8	6.1	6.2	0.16	0/5
42	IP 5931	G-227	63	6.6	6.3	5.8	5.9	5.8	6.1	0.24	0/5
43	IP 17493	G-212	65	7.5	6.7	5.4	5.4	5.1	6.0	0.66	1/5
	ICMR 11019 (Res. check)	G-249	53	6.2	6.3	5.9	5.3	5.8	5.9	0.26	0/5
	ICMB 95444 (Sus. check)	G-251	54	0.0	0.0	0.0	0.0	0.0	0.0	-	0/5
Environmental mean			59	4.6	4.3	4.5	4.5	4.5	4.5	-	-

(Continued)

TABLE 2 | Continued

Sl.no.	Entry	Code	DFF	FB scores at individual environments					Overall mean	Stability index (ASV)**	Frequency in top 10 resistant line group
				ND-17	ND-18	ND-19	BWL-19	JPR-19			
CV%			2.7	23.8	19.1	21.6	10.7	12.7	18.5	-	-
LSD			3.5	2.2	1.8	1.9	1.0	1.2	1.1	-	-

*Yellow color represents genotypes whose foliar blast score is above qualifying check. *Blue color represents qualifying checks under each environment. **Stability index worked out as described by Purchase et al. (2000); ASV = AMMI stability value (stable genotypes will have lower values). CV, coefficient of variation; LSD, least square difference; DFF, days to 50% flowering; NG, seeds not germinated; ND-17, New Delhi-2017; ND-18, New Delhi-2018; ND-19, New Delhi-2019; BWL-19, Bawal-2019; JPR-19, Jaipur-2019.

The phenotypic observations Y_{ijkm} on genotypes m in replicate j of block k of environment i was modeled as follows:

$Y_{ijkm} = \mu + e_i + r_{ij} + b_{ijk} + g_m + (ge)_{im} + (eg)_{jm} + \varepsilon_{ijkm}$
 where μ is the grand mean; e_i is the fixed effect of environment i ; g_m is the random effect of genotype m and is $\sim \text{NID}(0, \sigma_g^2)$; r_{ij} is the fixed effect of replication in the environment i ; b_{ijk} is the random effect of block k nested with replication j in the environment i and is $\sim \text{NID}(0, \sigma_b^2)$; $(ge)_{im}$ is the random effect of the interaction between genotype m and environment i and is $\sim \text{NID}(0, \sigma_{ge}^2)$; ε_{ijkm} is the random effect of the error variances.

Broad-sense heritability (H^2) for the traits in each environment and over combined environments was estimated from the variance components. For each environment, H^2 was calculated as $H^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_e^2/r)$; and for combined environments, H^2 was used as a measure of the trial's reliability in genotype evaluation in this study, with $H^2 = 0$ indicating that variations in genotypic mean in the trial are entirely attributable to random error and $H^2 = 1$ indicating that differences are entirely due to genetic effects (Yan and Holland, 2010).

The REML model also produced the best linear unbiased predictors (BLUPs) of each genotype, thereby adjusting the influence of the neighboring rows. These BLUPs were used for downstream analysis.

Heritability-Adjusted Genotype Plus Genotype × Environment Interaction Biplot Analysis

Best linear unbiased predictors values of the transformed disease mean score were stored in a 250 genotypes × five environments matrix M . The matrix was checked for missing data arising due to non-germination of seed in the individual environment and was corrected through imputation using the expectation-maximization algorithm implemented by R package, *bbplot/R Bilinear* as suggested by Gauch and Zobel (1990). Furthermore, heritability-adjusted scaling (Yan and Holland, 2010) was performed in R. The entries that were identified as resistant were further highlighted in GGE biplot construction for better visualization.

The GGE biplot was constructed by estimating each element of the matrix using the following formula, based on the first two principal components (PCs) resulting from singular value

decomposition (SVD) (Yan et al., 2000; Yan and Kang, 2003):

$$Y_{ij} = \mu + e_j + \sum_{n=1}^N \lambda_n \gamma_{in} \delta_{jn} + \varepsilon_{ij}$$

Where,

Y_{ij} = mean response of i^{th} genotype ($i = 1, \dots, i$) in the j^{th} environment ($j = 1, \dots, j$);

μ = grand mean;

e_j = environment deviations from the grand mean;

λ_n = the eigenvalue of PC analysis axis;

γ_{in} and δ_{jn} = genotype and environment PC scores for axis n ;

N = number of PCs retained in the model;

ε_{ij} = residual effect $\sim N(0, \sigma_e^2)$.

An “average environment coordination” (AEC) view of the GGE biplot was used to appraise genotypic response and stability. It facilitated genotype comparisons based on disease score mean and stability across environments within a “mega-environment” (Yan, 2001, 2002). The axis of the “AEC abscissa,” denoted by a single arrowed line, indicated higher mean performance of genotypes in terms of higher FB resistance, whereas the “AEC ordinate,” denoted by a line perpendicular to the AEC abscissa and passing through the origin of the biplot, represented genotype stability. Stability is represented by projections on the AEC abscissa connecting individual genotypes (Yan and Falk, 2002). Similarly, the “discriminating power vs. representativeness” view of the GGE biplot was constructed for the evaluation of test environments, where the “ideal” test environment should be both discriminating with respect to genotypes and representative of the “mega-environment” (Yan et al., 2007). In addition, a “desirability index” of the test locations has been compiled taking into account the relationship between the test environments and distance from the ideal genotype (Yan and Holland, 2010). Angles between the various environment vectors were used to judge the correlation among the environments in order to determine the relationship between test locations (Yan and Kang, 2003). Furthermore, a “which-won-where” view of the GGE biplot has been prepared to determine the superiority of the genotypes in different test environments as well as grouping of test environments into different “mega-environments” (Yan and Rajcan, 2002). Finally, bootstrapping (re-sampling process is repeated 10,000 times to

TABLE 3 | Analysis of variance for foliar blast score and days to 50% flowering at individual environment under rainfed conditions.

Traits	Effects	Source of variance	DF	New Delhi-17				New Delhi-18				Bawal-19				Jaipur-19			
				Variance	SE	F value ^a and LRT value ^b	p values	Variance	SE	F value ^a and LRT value ^b	p values	Variance	SE	F value ^a and LRT value ^b	p values	Variance	SE	F value ^a and LRT value ^b	p values
Foliar blast score	Fixed	Replication	1	–	–	0.10	0.758	–	–	0.24	0.628	–	–	–	0.810	–	–	–	0.597
	Random	Genotype	249	2.89	1.7	165.70	<0.001	3.616	1.9	263.12	<0.001	3.46	1.86	430.44	<0.001	3.40	1.85	230.45	<0.001
		Block (Replication)	48	0.06	0.25	1.55	0.213	0.011	0.11	0.11	0.743	0.04	0.19	4.56	0.033	0.02	0.12	0.14	0.707
Days to 50% flowering	Fixed	Replication	1	–	–	0.34	0.572	–	–	0.036	0.852	–	–	–	0.255	–	–	–	0.925
	Random	Genotype	249	47.86	6.9	717.77	<0.001	34.728	5.893	393.480	<0.001	62.79	7.92	570.86	<0.001	68.37	8.27	1,188.93	<0.001
		Block (Replication)	48	0.00	0.0	0.00	1.000	0.032	0.180	0.032	0.858	0.25	0.50	2.57	0.109	0.00	0.00	0.00	1.000
		Residuals	201	1.30	1.1	–	<0.001	4.179	2.044	0.943	<0.001	3.19	1.79	0.975	<0.001	0.28	0.53	0.978	<0.001
		Heritability				0.987													

DF, degree of freedom; SE, standard error. ^aConcerning the fixed effect components. ^bConcerning the random effect components.

obtain 10,000 bootstrap samples) is used to assess the validity of GGE biplot. These bootstrap samples were later used to construct CL at the 95% level for individual genotype and environment PC scores as suggested by Hu and Yang (2013). The raw data with columns representing the environments ($p = 5$) and rows representing the genotypes ($n = 43$) were later average-centered for each environment, resulting in a mean of zero for each of the p dimensions of raw data. The data were subjected to bidirectional bootstrapping and procrustes rotation in the R package “bbplot/R,” and the confidence regions were computed using a distribution-free approach implemented in the R package “distfree.cr/R” based on the empirical distribution of the aligned genotypic or environmental scores from all bootstrap samples.

RESULTS

The diversity panel of 250 accessions along with checks was subjected to phenotyping for FB incidence and DFF at five different environments. Phenotypic data collected from the population at three different locations during the rainfed seasons of 2017, 2018, and 2019 were statistically analyzed to determine variance components. Hereafter, the five environments are denoted as ND-17 (New Delhi-17), ND-18 (New Delhi-18), ND-19 (New Delhi-19), BWL-19 (Bawal-19), and JPR-19 (Jaipur-19).

Identification of Climatic Factors Influencing Foliar Blast Infection

Weather parameters observed in each environment during tillering to hard dough stage of plant growth (30–60 DAS) are presented in Table 1, and their influence over FB score is elucidated in the CCA diagram (Figure 2). CCA biplot explained 80.77% of the total variation between the site weather parameters and the FB score. The first CCA axis explained 48.08%, and the second CCA axis explained 32.69% of the total variation. Maximum temperature (Max. Temp), minimum temperature (Min. Temp), and maximum humidity (Max.RH) were the main determinants and were positively associated with the increase of FB score during the critical growth period of the crop. Cumulative rainfall recorded during the critical period of growth of the crop was not associated with FB score. Based on spatial angular proximity of identified climatic determinants (Max. Temp, Min. Temp, and Max. RH), JPR-19 was found to be conducive for FB disease in pearl millet.

Mean Performance of Genotypes and Analysis of Variance

The experiment was executed systematically in all the five environments. Susceptible check, ICMB 95444, included in the experiment exhibited uniform FB-susceptible reaction indicating the availability of sufficient inoculums for disease screening. The mean value, standard deviations, and the frequency distribution for FB scores within and across years and locations indicate

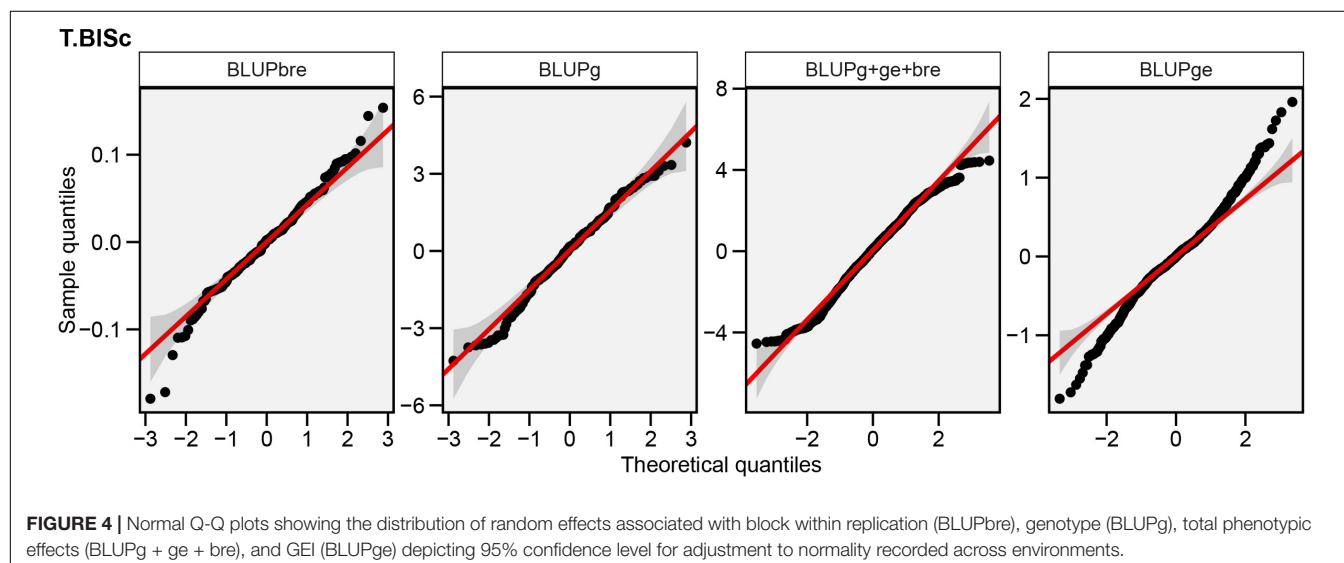
TABLE 4 | Combined analysis of variance for foliar blast resistance and days to 50% flowering in pearl millet.

Traits	Effects	Source of variance	DF	Variance	SE	F value ^a and LRT value ^b	p values	H ²
Foliar blast score	Fixed	Environments	4			6.017	<0.0001	
		Environments (Replication)	5			0.191	0.966	
	Random	Genotype	249	2.792	1.671	1,106.751	<0.0001	0.887
		Genotype × Environments	996	0.474	0.689	192.564	<0.0001	
		Environments (Replication × Block)	240	0.015	0.124	2.057	0.151	
Days to 50% flowering	Fixed	Residuals	1,245	0.709	0.842			
		Environments	4			179.725	<0.0001	
	Random	Environments (Replication)	5			0.256	0.937	
		Genotype	249	38.105	6.173	811.343	<0.0001	0.968
		Genotype × Environments	996	16.438	4.054	1,610.839	<0.0001	
		Environments (Replication × Block)	240	0.093	0.305	3.160	0.075	
		Residuals	1,245	2.521	1.588			

DF, degree of freedom; SE, standard error; H², heritability. ^aConcerning the fixed effect components. ^bConcerning random effect components.

that the lines exhibited similar levels of disease severity in all five environments (**Figure 3A**) although slight differences in distribution are evident from the histogram for ND-17. The average FB score in ND-17 was marginally higher than the average scores, but the distribution pattern was flatter, indicating a higher level of variability (SD = 1.7). ND-18 showed the highest variability (SD = 1.88), yet the distribution was slightly skewed toward the resistant side, with FB score considerably lower than the remaining four environments. Moreover, there was a presence of crossover G × E interaction among genotypes for FB scores, which were evident from the heatmap visualization of the GE interaction (**Figure 3B**). From the pooled data, out of 250 genotypes, none of the genotypes showed a highly resistant reaction (score = 9) to FB. However, at the hard dough

stage, 43 genotypes were reported as resistant (R), 118 as moderately resistant/susceptible (MR/MS), 70 as susceptible (S), and 19 as highly susceptible (HS) to FB (**Figure 3C**). Out of 43 resistant lines, five were late in flowering (51–54 days), whereas 38 lines flowered very late (>54 days). Among 43 resistant lines, 26 genotypes exhibited a resistance reaction above the qualifying check, ICMR 11009 (score = 6.5), across environments (**Table 2**). Also, five more genotypes have an FB score that is at par with the qualifying check. Genotypes IP 11353, IP 22423, IP 7941, IP 7910, IP 12322, and IP 3106 were consistently showing higher resistance and outperformed the corresponding qualifying checks identified for each environment (indicated in blue color). Considering the frequency of appearance in the top 10 lines based on FB scores, only three genotypes, namely, IP 11353 (G-62), IP



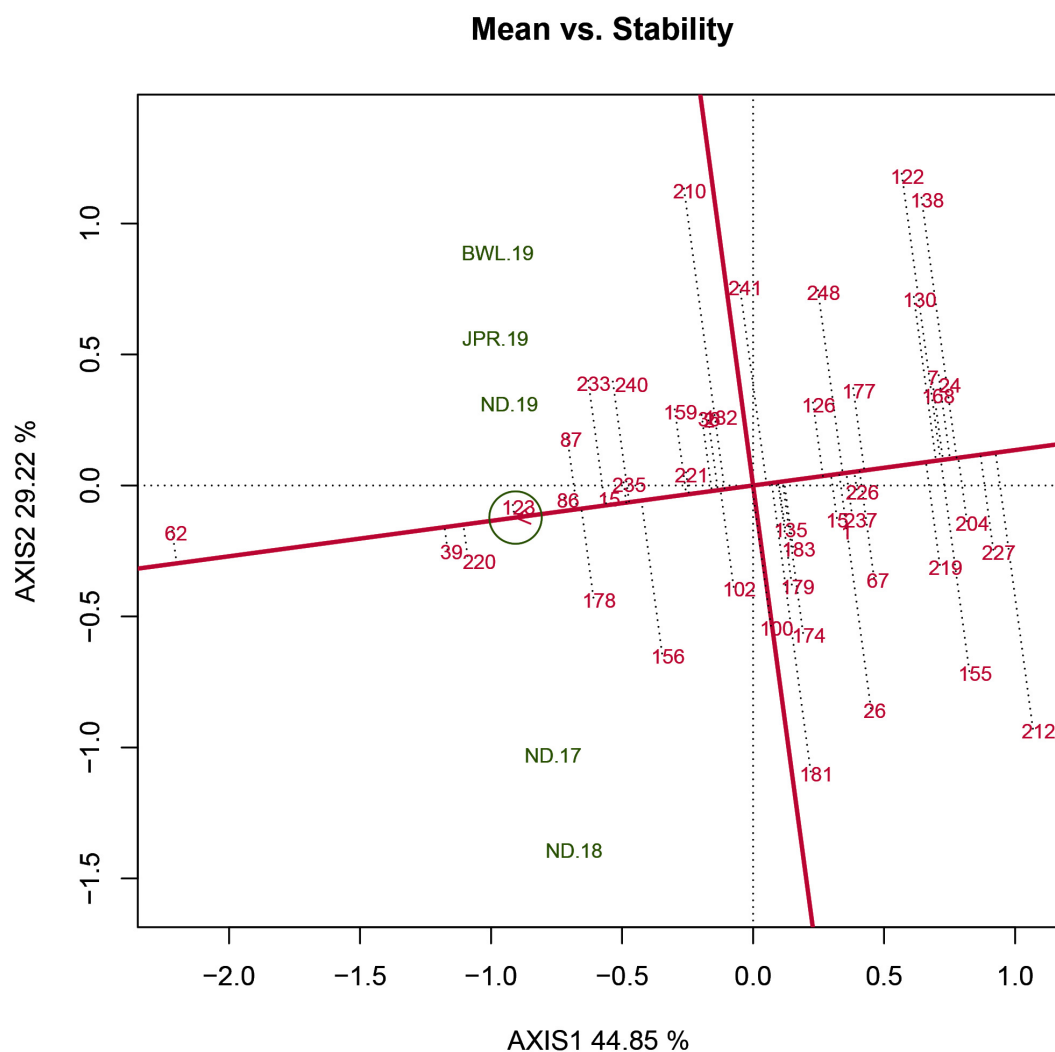


FIGURE 5 | Mean vs stability view of the GGE biplot of 43 resistant genotypes across five testing environments. There was heritability-adjusted scaling of data where the environment standardized data were multiplied by the heritability in each environment (transform = HA), and data were centered by means of the environments (centering = 2). The biplot was based on “row metric preserving” [singular value partition (SVP) = 1], which means that the singular values were partitioned into the genotype eigenvectors for visualizing the correlation among genotypes. Numbers correspond to genotypes as listed in **Supplementary Table 1**. Environment: ND-17, New Delhi-2017; ND-18, New Delhi-2018; ND-19, New Delhi-2019; BWL-19, Bawal-2019; JPR-19, Jaipur-2019.

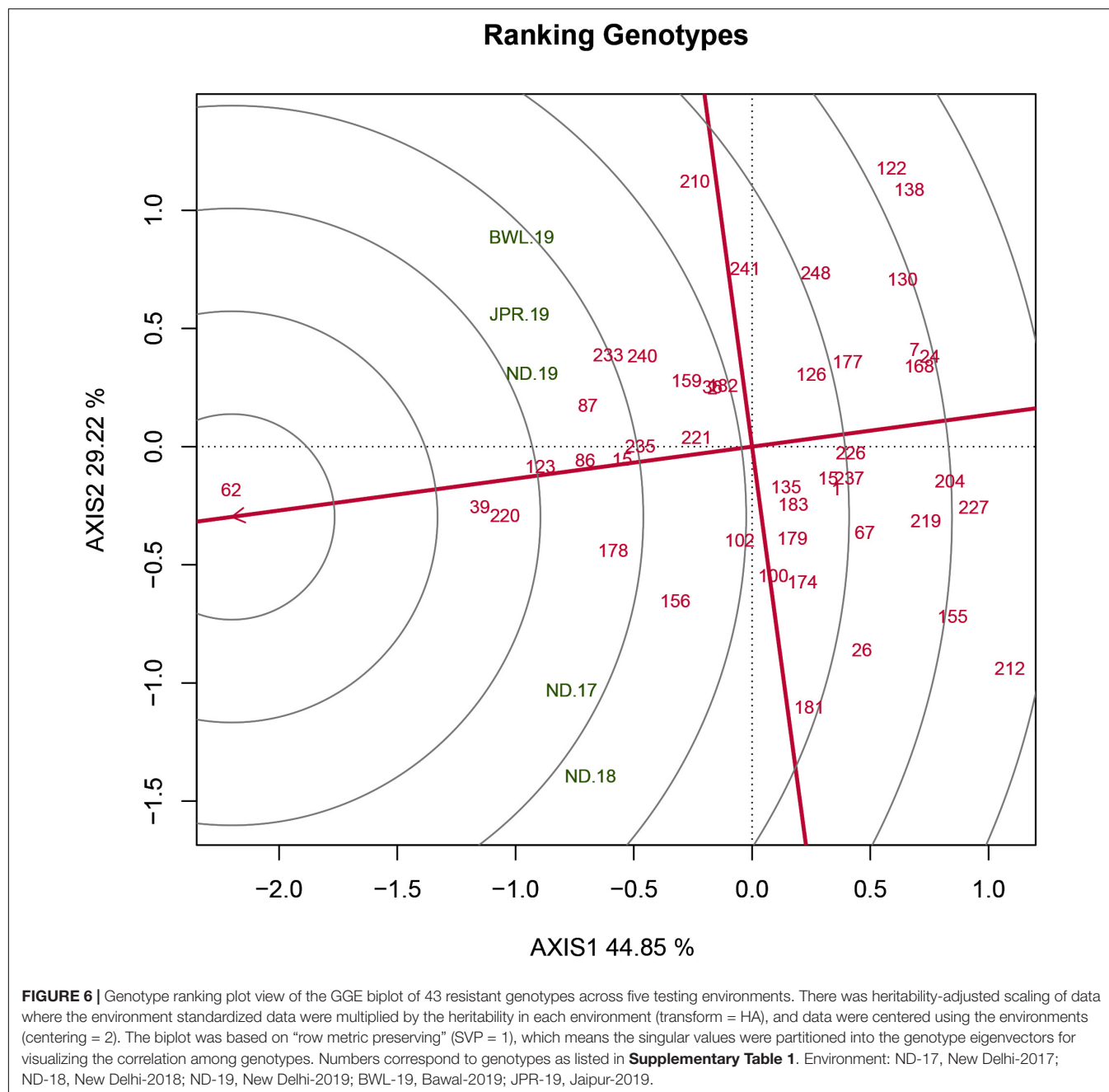
22423 (G-39), and IP 7941 (G-220), were found to be more consistent with the lowest stability index below 0.1 in all the five test environment.

Analysis of Variance and Estimation of Heritability

Analyses of variance of individual environment indicated significant genotypic differences ($p < 0.001$) for DFF and FB disease score at the hard dough stage. FB disease score has shown the highest genotypic variance at ND-18, and DFF has shown the highest variance at BWL-19 (Table 3). Combined ANOVA also revealed significant genotypic differences along with significant environment and GEI effects ($p < 0.001$) for DFF and FB disease score at the hard dough stage. The contribution of genotypic variance

toward total phenotypic variance was higher for FB disease score, whereas the contribution of GEI variances was higher compared with genotypic variance for DFF (Tables 3, 4). The contribution of environment variance was very low for both DFF and FB disease scores. Probability plots of residuals *versus* expected values indicated no discernible trend, implying that the assumptions of independence and equal variance were fulfilled.

In the current analysis, both traits were strongly heritable (>0.60) in individual environments, as per the Robinson et al. (1966) scale of heritability (Table 3). Compared with FB disease score, DFF was found to be more heritable. Broad-sense heritability for both the traits was also higher (>0.60) across five environments (Table 4). However, when assessed on the basis of pooled environment, a

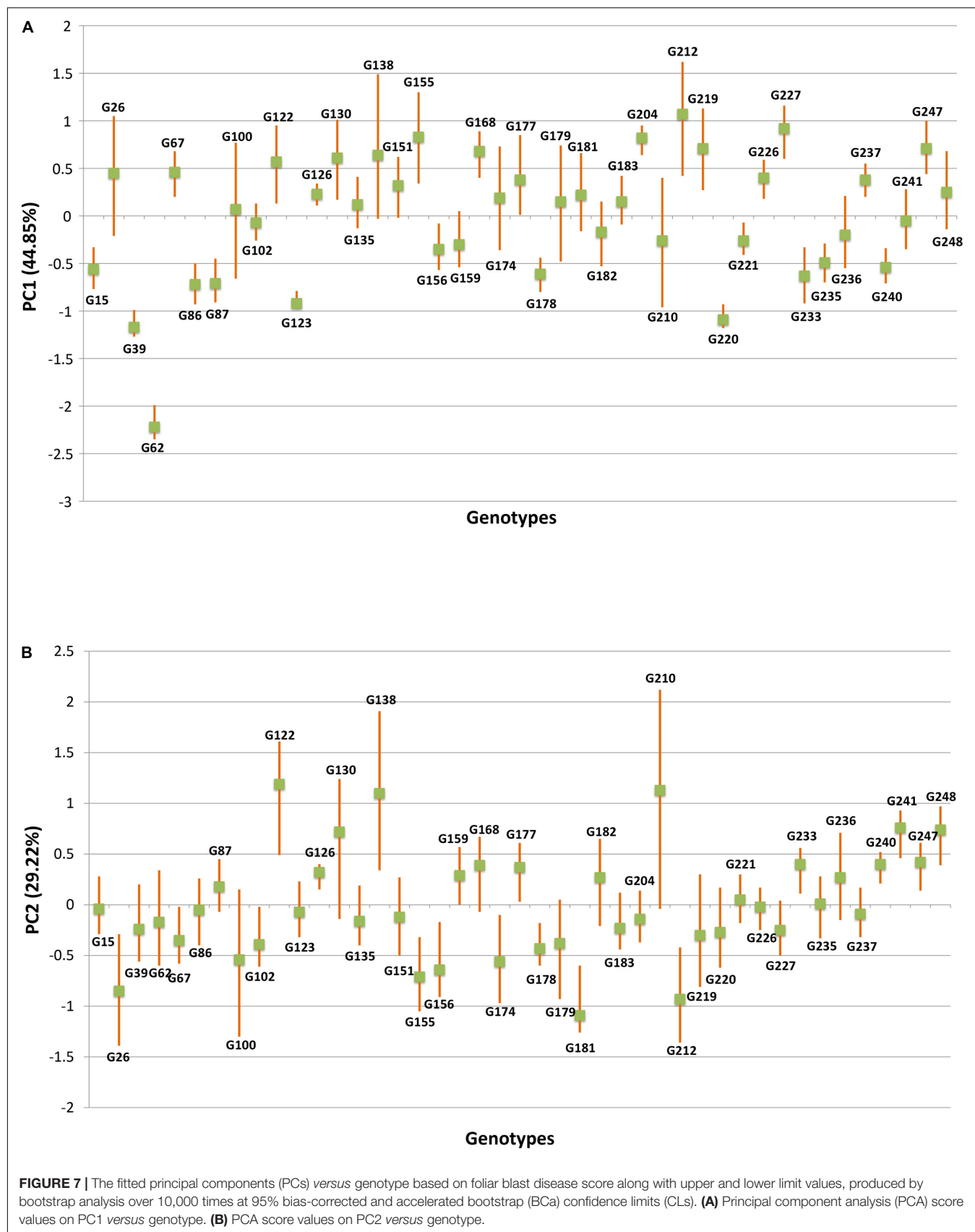


partitioning of GEI component lowered heritability for both the traits across environments. For foliar disease score, broad-sense heritability ranged from 0.83 (ND-17) to 0.96 (JPR-19); and in the pooled environment analysis, it was 0.89. Similarly, for DFF, it ranged from 0.94 (ND-18) to 0.99 (BWL-19) in case of the individual environment, while in case of pooled data, heritability estimate was 0.97. High heritability for both the traits indicated that the genotypic differences observed are mainly due to genetic effects. Also, there was a moderate, significant positive correlation between FB disease score (transformed data) and DFF ($r = 0.503$, $p < 0.0001$), indicating that FB

disease resistance in pearl millet is associated with very late flowering (**Figure 3C**).

Estimation of Best Linear Unbiased Predictor Values and Imputation of Missing Data

Predicted genetic values (BLUP) were estimated to guide the inferences based on a multi-environment GGE model with reduced biases arising from uncontrollable factors (**Figure 4**). It was useful for recommending genotypes with the minimum likelihood of error while recommending



them for a specific environment. For the individual environments, the predicted FB score ranged from 0.81 to 8.21 in ND-17, from 0.48 to 8.56 in ND-18, from 0.19 to 8.79 in ND-19, from 0.61 to 8.50 in BWL-19, and from 0.29 to 8.84 in JPR-19. Under all the environments, genotype G-13 (IP 4020) had the lowest resistance score, and genotype G-62 (IP 11353) had the highest level of FB resistance. Similarly, across the environment, the estimate of genotype (random effect) for FB score ranged from 0.28 (G-13) to 8.77 (G-62). Prior to heritability scaling, the genotype–environment BLUP matrix (M) was analyzed for missing data. It was observed that around 1% of missing data were solved by imputation by the expectation–maximization algorithm.

Detection of Ideal Genotype Based on Mean Versus Stability

The HA-GGE biplot is the most precise way to detect ideal genotypes. An ideal genotype should have both the highest mean performance and the lowest interactions with the environment. The mean *vs* stability biplot view of the HA-GGE biplot was generated based on the principle of genotype-focused singular value partition (SVP) ($SVP = 1$) as suggested by Yan (2002). This biplot view portrays the ranking of genotypes based on their average FB score across environments (Figure 5). PC 1 and PC 2 explained 44.85 and 29.22%, respectively, of the total variation of the environment scoring. The single arrowhead line passing through the origin, the AEC abscissa, indicated a highly resistant genotype with a lower FB score. Therefore, genotypes positioned downstream of the arrow are considered promising for FB resistance reaction. IP 11353 (G-62), IP 22423 (G-39), IP 7910 (G-123), and IP 7941 (G-220) were positioned downstream of the biplot origin and, therefore, experienced less FB score. Moreover, the stability of the genotype could be accessed through the length of the projection in both directions from the AEC abscissa, that is, the AEC ordinate. Thus, if the genotype had greater projection from the AEC abscissa, it would be less stable. Considering both mean performance and stability, IP 11353 (G-62) was the ideal genotype, having less disease score and high stability (Figure 6). Since the distance between two genotypes should always be estimated by Euclidian distance, genotypes that are closer to the ideal genotype are considered to be desirable (Yan and Tinker, 2006). Therefore, IP 22423 (G-39), IP 7910 (G-123), and IP 7941 (G-220) were identified as desirable genotypes with lesser FB scores and almost consistent performance. Furthermore, using the CL at the 95% level for individual genotypic scores on FB as well as environmental scores corresponding to PC 1 and PC 2 (Supplementary Table 2), bootstrapping revealed that PC 1 contributed more to the significant differences among genotypes, as seen on the biplot (Figure 7). In terms of FB scores, it was established that the ideal genotype IP 11353 (G-62) was statistically different from the three desirable genotypes, IP 22423 (G-39), IP 7910 (G-123), and IP 7941, based on PC 1 scores (lower limit, 2.35; and upper limit, 1.99) (G-220). Three desirable genotypes, on the other hand,

TABLE 5 | Standardized test environment evaluation parameters.

Environment	Discriminative	Representative	Desirability	Desirability
		-ness	index	index rank
ND-17	1.257	0.771	0.969	4
ND-18	1.446	0.629	0.910	5
ND-19	1.100	0.932	1.025	2
BWL-19	1.361	0.750	1.021	3
JPR-19	1.187	0.864	1.026	1

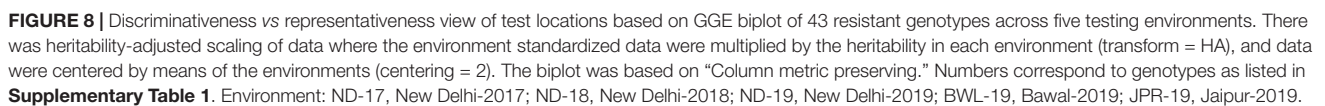
showed no significant differences in their PC 1 scores for both parameters.

Evaluation of Testing Locations Based on Discrimination Power Versus Representativeness and Desirability Index

Three parameters, namely, discrimination power (ability to segregate the tested genotypes), representativeness (ability to represent the ME) and desirability index (the joint response of both discriminating power and representativeness) are crucial for assessing the test environment in the GGE biplot approach. In the HA-GGE biplot, the length of the environmental vector, which is approximately the square root of heritability, represents discrimination ability. The angle between the environmental vectors and the AEC abscissa specifies the representativeness of the testing location. The environment becomes more representative when the angle among the test environment with AEC abscissa becomes more acute. In analysis, it was found that among the test environments, BWL-19 had the longest environmental vectors, rendering it as the most “discriminating location” with the ability to discriminate genotypes from other sites. However, in the case of representativeness, ND-19 showed a minimum angle with average environment followed by JPR-19. Hence, the desirability index was worked out to identify the most ideal testing location accounting for both discrimination ability and representativeness (Table 5). The Jaipur center having the highest desirability index (1.026) was identified as an ideal or a type I testing location for testing of a mini core collection or advance breeding materials against FB disease. Since ND-19 had also been included in the same sector, therefore, it can be considered as a supplementary or type II location for testing pearl millet genotypes against FB (Figure 8).

Relationship Among Environments and Mega-Environment Delineation

In the current study, “which-won-where” biplot for FB score created a hexagon with six genotypes, G-62, G-210, G-122, G-138, G-242, and G-181, at the vertices (Figure 9). The equality lines divided the polygon into six sectors effectively. Five testing environments were spread in two sectors within the biplot: three in one and two in another. This illustrated that the testing locations could be divided into MEs. The first ME (ME-I) was represented by ND-19, JPR-19, and BWL-19, with IP 11353 (G-62) having the highest FB resistance as the winning genotypes.



In the present study, genotypes showed significant differences for both DFF and FB incidence in all five environments. Even though mean FB incidence was almost similar for all the five environments, ND-17 showed a slightly higher incidence followed by ND-19 and BWL-19. With respect to environment variables, it was observed that maximum temperature (Max. Temp), minimum temperature (Min. Temp), and maximum relative humidity (Max. RH) were slightly higher at ND-17

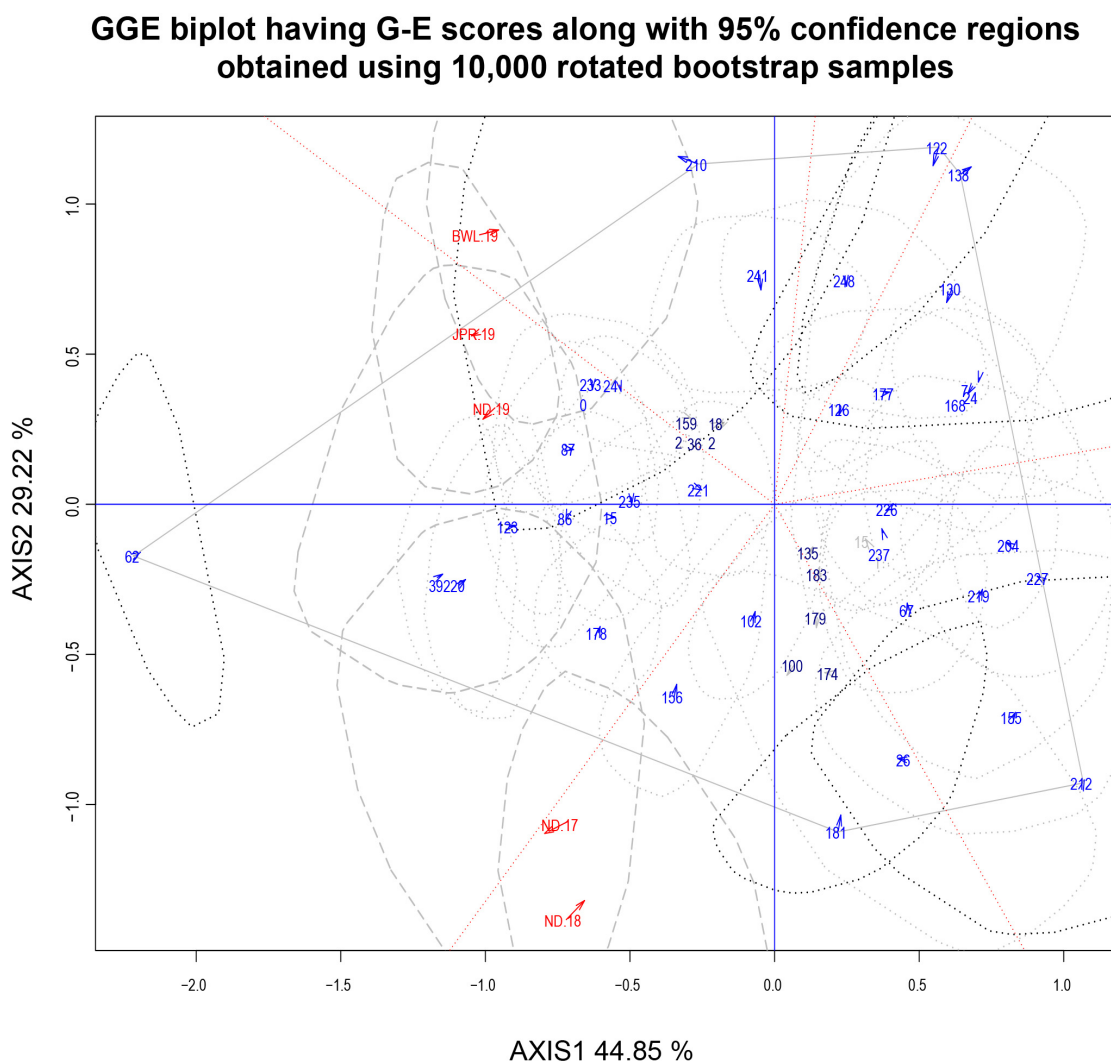


FIGURE 9 | Which-won-where biplot view of 43 genotypic scores and five environmental scores constructed along with the 95% confidence regions using 10,000 rotated bootstrap samples. There was heritability-adjusted scaling of data where the environment standardized data were multiplied by the heritability in each environment (transform = HA), and data were centered by means of the environments (centering = 2). The biplot was based on “row metric preserving” [singular value partition (SVP) = 1], which means that the singular values were partitioned into the genotype eigenvectors for visualizing the correlation among genotypes. Numbers correspond to genotypes as listed in **Supplementary Table 1**. Environment: ND-17, New Delhi-2017; ND-18, New Delhi-2018; ND-19, New Delhi-2019; BWL-19, Bawal-2019; JPR-19, Jaipur-2019.

followed by ND-19. These factors might have played a significant role in determining the emergence of the blast disease as also reported by Pattanayak and Das (2020).

The G, E, and $G \times E$ interactions displayed significant differences as revealed by ANOVA. The presence of a significant $G \times E$ interaction indicated that the FB incidence of tested genotypes varied across environments, which could be attributed to different agro-ecologies with varying longitude, latitude, and elevation. Significant GEI also suggested the need to develop FB-resistant lines with specific adaptation to target ecology. Furthermore, genotypic variance contributed more to disease resistance than the $G \times E$ relationship, suggesting that genetic variation accounted for the most of the variation in disease reactions. Persaud and Saravanakumar (2018) also reported

greater contribution of genotypic factor over $G \times E$ interaction factor while conducting a multilocation experiment in case of rice FB in which 76.02% of the total SS was attributed to genotype (G) effect, 3.10% to environment (E) effects, and 20.88% to GEI effects. Similar results were also reported by Beyene et al. (2012) in maize foliar disease resistance, Sharma et al. (2012) in chickpea wilt incidence, and Sharma et al. (2016) in pigeon pea-*Fusarium udum* interaction.

An initial study of 250 genotypes facilitated the selection of 43 resistant genotypes for HA-GGE biplot analysis. The complex GEIs were simplified in different PCs and graphically presented against various PCs in GGE biplot analysis, and their contribution justified the GGE biplot's utility in explaining sources of variation (Yan and Tinker, 2005). In the present

study, the first two PCs clarified more than 70% of the total variance, indicating that the variability for FB resistance reaction is adequately represented. The “mean vs. stability view” of the biplot for the trait represented differential responses of tested genotypes to diverse environments due to the existence of crossover interactions. Genotype ranking in terms of resistance to blast was observed to change from one environment to another. The genotypes IP 11353 (G-62), IP 22423 (G-39), IP 7910 (G-123), and IP 7941 (G-220) were positioned downstream of the biplot origin and, therefore, experienced less FB score and were considered resistant. Among these genotypes, IP 11353 (G-62) was considered to be the ideal genotype owing to its higher disease resistance and smaller interaction with the environment in the form of a high projection from the AEC abscissa (Yan and Falk, 2002). Genotypes that are in proximity with “ideal” were considered “desirable” due to their high genetic relationship with the “ideal” genotype (Yan and Tinker, 2005). IP 22423 (G-39), IP 7910 (G-123), and IP 7941 (G-220) were identified as desirable genotypes owing to their proximity to the ideal genotypes that differ in their ability to respond to fungal infection by inducing long-lasting, broad-spectrum, and systemic resistance. Bootstrapping at 95% CL improved the precision of the visual observation recorded on promising genotypes. The ideal genotype revealed a significant statistical difference from the desirable genotypes. However, all of the desirable genotypes were overlapping. HA-GGE biplot has successfully detected stable and resistant genotypes in various crops (Silva et al., 2011; Sillero et al., 2017; Sánchez-Martín et al., 2017; Parihar et al., 2018; Singh et al., 2020). Thus, the “ideal” genotypes, along with any one of the “desirable” genotypes having durable resistance, would be valuable genetic resources in the upcoming comprehensive resistance breeding program of pearl millet.

Heritability-adjusted genotype plus genotype \times environment interaction biplot identified the superior testing location, facilitating complete resource allocation with minimum trial cost without compromising trial heritability and genetic gain under selection (Yan, 2001; Yan and Tinker, 2006; Yan and Holland, 2010). Previous reports also claimed that, assuming adequate discriminating capacity, “representativeness” is the most important factor to be considered when deciding how a test location should be used in genotype evaluation (Yan et al., 2007). The square root of heritability ($\sqrt{H^2}$) of each test environment based on vector length and the representativeness as its genetic correlation with other test environments (r) based on the angle between two test environments can be assessed by HA-GGE biplots (Allen et al., 1978; Flores et al., 2013). Considering both the parameters, Jaipur center (JPR-19), with the highest desirability index, was recognized as the ideal or type I testing location for testing advanced breeding materials for FB-resistant progenies during the early breeding stage. The existence of non-crossover GEI (consistent performance of genotype) suggested the presence of a close relationship among the test locations. Thus, HA-GGE biplot is the most precise method for proper delineation of an ideal testing location.

The only way to accomplish consistent genotype performance within a given sector is to divide testing locations into distinct “mega-environments.” “A mega-environment can be described

as a group of analogous locations delivering similar genotypic responses when sharing the same set of genotypes across the year” (Yan and Rajcan, 2002). The “mega-environment” can be effectively depicted using a “which-won-where” view of GGE biplot methodology (Gauch and Zobel, 1997; Yan and Kang, 2003; Yan et al., 2007). The aim of ME diagnosis is to better understand the complex GEI pattern that occurs within that region in order to exploit specific adaptations and maximize selection responses (Yan, 2011). In the current investigation, the HA-GGE biplot was able to separate all of the testing locations into two distinct MEs to aid the restructuring of agro-ecological zones. Year 2019 was separated as single ME (ME-I) in which the most desirable environment JPR-19 was also included. Environmental conditions of the Jaipur location were found to be more conducive for FB incidence. Hence, Jaipur location owing to its informative role in the present study can be selected in future FB testing programs. Bootstrapping at 95% CL improved the precision of recommendation of a testing location and ME delineation. However, reported groupings of environments need to be further reconfirmed using multi-location testing data over more number of years as also reported by earlier authors (Yan et al., 2000; Silva et al., 2011; Phuke et al., 2017; Sánchez-Martín et al., 2017; Das et al., 2020). Also, the performance and stability of all selected materials advocate additional testing in central and southern pearl millet-growing regions of the country for the future development of elite FB-resistant cultivars.

CONCLUSION

In the present study, genotypic effects and GEI exhibited the greatest effect in comparison with the environment alone for FB resistance in pearl millet. Based on the HA-GGE biplot, all of the tested locations could be grouped into two distinct MEs with winning genotypes. It confirms the presence of crossover-type GEI and emphasizes breeding for environment-specific adaptability. More importantly, among the tested genotypes, IP 11353 was recognized as “ideal,” and IP 22423, IP 7910, and IP 7941 were recognized as “desirable” genotypes, having stable resistance against the disease. Salient findings obtained in the present study were also validated by bootstrapping at 95% CL. This study was also able to reorganize delineated MEs and advocates precise testing of materials with optimization of resources in future breeding programs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

RS and SPS designed and supervised the overall research and contributed to the preparation of the manuscript. CS provided technical guidance and liaised among multi-environment

locations. SMS, YY, and LS executed the field experiments. SMS and GP performed the phenotyping and disease scoring. AR and SLS carried out statistical analysis and preparation of the draft manuscript. SLS and NS performed manuscript review. RS edited the manuscript for final submission. All authors contributed to the article and approved the submitted version.

FUNDING

SMS is also extremely thankful for the funding provided by the ICAR-World Bank funded by National Agricultural Higher Education Project in order to conduct this research.

REFERENCES

- Adhikari, S., Joshi, S. M., Athoni, B. K., Patil, P. V., and Jogaiah, S. (2020). Elucidation of genetic relatedness of *Magnaporthe grisea*, an incident of pearl millet blast disease by molecular markers associated with virulence of host differential cultivars. *Microb. Pathog.* 149:104533. doi: 10.1016/j.micpath.2020.104533
- AICPMIP (2020). *Annual Report, All India Co-Ordinated Pearl Millet Improvement Project*. Mandor: AICPMIP, Indian Council of Agricultural Research.
- Allen, F. L., Comstock, R. E., and Rasmusson, D. C. (1978). Optimal environments for yield testing. *Crop Sci.* 18, 747–751. doi: 10.2135/cropsci1978.0011183x001800050013x
- Anuradha, N., Satyavathi, C. T., Bharadwaj, C., Nepolean, T., Sankar, S. M., Singh, S. P., et al. (2017). Deciphering genomic regions for high grain iron and zinc content using association mapping in pearl millet. *Front. Plant Sci.* 8:412. doi: 10.3389/fpls.2017.00412
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed effects models using lme4. *J. Stat. Softw.* 67, 1–48.
- Beyene, Y., Mugo, S., Tefera, T., Gethi, J., Gakunga, J., Ajanga, S., et al. (2012). Yield stability of stem borer resistant maize hybrids evaluated in regional trials in East Africa. *Afr. J. Plant Sci.* 6, 77–83.
- Das, A., Gupta, S., Parihar, A. K., Singh, D., Chand, R., Pratap, A., et al. (2020). Delineating genotype \times environment interactions towards durable resistance in mung bean against *Cercospora* leaf spot (*Cercospora canescens*) using GGE biplot. *Plant Breed.* 139, 639–650. doi: 10.1111/pbr.12789
- Flores, F., Hybl, M., Knudsen, J. C., Marget, P., Muel, F., Nadal, S., et al. (2013). Adaptation of spring faba bean types across European climates. *Field Crops Res.* 145, 1–9. doi: 10.1016/j.fcr.2013.01.022
- Gauch, H. G. Jr., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.* 37, 311–326. doi: 10.2135/cropsci1997.0011183x003700020002x
- Gauch, H. G., and Zobel, R. W. (1990). Imputing missing yield trial data. *Theor. Appl. Genet.* 79, 753–761. doi: 10.1007/bf00224240
- Goud, T. Y., Sharma, R., Gupta, S. K., Devi, G. U., Gate, V. L., and Boratkar, M. (2016). Evaluation of designated hybrid seed parents of pearl millet for blast resistance. *Ind. J. Plant Prot.* 44, 83–87.
- Govindaraj, M., Yadav, O. P., Rajpurohit, B. S., Kanatti, A., Rai, K. N., and Dwivedi, S. L. (2020). Genetic Variability, Diversity and Interrelationship for twelve grain minerals in 122 commercial pearl millet cultivars in India. *Agric. Res.* 9, 516–525. doi: 10.1007/s40003-020-00470-7
- Gupta, S. K., Sharma, R., Rai, K. N., and Thakur, R. P. (2012). Inheritance of foliar blast resistance in pearl millet (*Pennisetum glaucum*). *Plant Breed.* 131, 217–219. doi: 10.1111/j.1439-0523.2011.01929.x
- Hu, Z., and Yang, R. C. (2013). Improved statistical inference for graphical description and interpretation of genotype \times environment interaction. *Crop Sci.* 53, 2400–2410. doi: 10.2135/cropsci2013.04.0218
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2015). *lmerTest: Tests in Linear Mixed Effects Models. R Package Version 2.0-25*. Available online at: <http://CRAN.R-project.org/package=lmerTest> (accessed August 2020).

ACKNOWLEDGMENTS

The authors are grateful to the A. K. Singh (Director, IARI, New Delhi) Vinod (Professor, Division of Genetics, IARI, New Delhi) and the PG School, ICAR-IARI, for providing the necessary facilities for the streamlined conduct of the research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.656158/full#supplementary-material>

- Nayaka, C. S., Srivastava, R. K., Udayashankar, A. C., Lavanya, S. N., Prakash, G., Bishnoi, H. R., et al. (2017). *Magnaporthe Blast of Pearl Millet in India – Present Status and Future Prospects*. Mandor: All India Coordinated Research Project on Pearl Millet (Indian Council of Agricultural Research), 51.
- Pankaj, P. K., Gaur, M. K., Nirmala, G., Maruthi, V., Samuel, J., and Reddy, K. S. (2020). “Diversification and land use management practices for food and nutritional security under the climate change scenario in arid and semi-arid regions of India,” in *Food Security and Land Use Change Under Conditions of Climatic Variability*, eds V. Squires, and M. Gaur (Cham: Springer), 281–309. doi: 10.1007/978-3-030-36762-6_15
- Parihar, A. K., Basandrai, A. K., Kushwaha, K. P. S., Chandra, S., Singh, K. D., Bal, R. S., et al. (2018). Targeting test environments and rust-resistant genotypes in lentils (*Lens culinaris*) by using heritability-adjusted biplot analysis. *Crop Pasture Sci.* 69, 1113–1125. doi: 10.1071/cp18259
- Pattanayak, S., and Das, S. (2020). Outbreak of rice blast on the coastal region of South-Eastern India. *Int. J. Agric. Environ. Biotechnol.* 13, 59–70.
- Patterson, H. D., and Thompson, R. (1971). Maximum likelihood estimation of variances. *Biometrika* 58, 545–554.
- Persaud, R., and Saravanakumar, D. (2018). Screening for blast resistance in rice using AMMI models to understand G \times E interaction in Guyana. *Phytoparasitica* 46, 551–568. doi: 10.1007/s12600-018-0681-7
- Phuke, R. M., Anuradha, K., Radhika, K., Jabeen, F., Anuradha, G., Ramesh, T., et al. (2017). Genetic variability, genotype \times environment interaction, correlation, and GGE Biplot analysis for grain iron and zinc concentration and other agronomic traits in RIL population of Sorghum (*Sorghum bicolor* L. Moench). *Front. Plant Sci.* 8:712. doi: 10.3389/fpls.2017.00712
- Prakash, G., Sriniwasa, N., Sankar, S. M., Singh, S. P., and Satyavathi, C. T. (2016). Standardization of pearl millet blast (*Magnaporthe grisea*) phenotyping under artificial conditions. *Ann. Agric. Res. Ser.* 37, 200–205.
- Purchase, J. L., Hatting, H., and van Deventer, C. S. (2000). Genotype \times environment interaction of winter wheat (*Triticum aestivum* L.) in South Africa: II. Stability analysis of yield performance. *S. Afr. J. Plant. Soil* 17, 101–107. doi: 10.1080/02571862.2000.10634878
- Robinson, H. F., Comstock, R. E., and Harvey, P. H. (1966). Estimation of heritability and the degree of dominance in corn. *Agron. J.* 41, 353–359. doi: 10.2134/agronj1949.00021962004100080005x
- Sánchez-Martín, J., Rispail, N., Flores, F., Emeran, A. A., Sillero, J. C., Rubiales, D., et al. (2017). Higher rust resistance and similar yield of oat landraces versus cultivars under high temperature and drought. *Agron. Sustain. Dev.* 37:3.
- Sharma, M., Babu, T. K., Gaur, P. M., Ghosh, R., Rameshwar, T., Chaudhary, R. G., et al. (2012). Identification and multi-environment validation of resistance to *Fusarium oxysporum* f. sp. ciceris in chickpea. *Field Crops Res.* 135, 82–88. doi: 10.1016/j.fcr.2012.07.004
- Sharma, M., Ghosh, R., Telangre, R., Rathore, A., Saifulla, M., Mahalinga, D. M., et al. (2016). Environmental influences on pigeonpea-*Fusarium udum* interactions and stability of genotypes to *Fusarium* wilt. *Front. Plant Sci.* 7:253. doi: 10.3389/fpls.2016.00253
- Sharma, R., Gate, L. V., and Madhavan, S. (2018). Evaluation of fungicides for the management of pearl millet [*Pennisetum glaucum* (L.)] blast caused by *Magnaporthe grisea*. *Crop Prot.* 112, 209–213. doi: 10.1016/j.cropro.2018.05.021

- Sharma, R., Goud, T. Y., Prasad, Y. P., Nimmala, N., Kadavani, D. L., Mathur, A. C., et al. (2021). Pathogenic variability amongst Indian isolates of *Magnaporthe grisea* causing blast in pearl millet. *Crop Prot.* 139:105372. doi: 10.1016/j.cropro.2020.105372
- Sharma, R., Sharma, S., and Gate, V. L. (2020). Tapping *Pennisetum violaceum*, a wild relative of pearl millet (*Pennisetum glaucum*), for resistance to blast (caused by *Magnaporthe grisea*) and rust (caused by *Puccinia substriata* var. indica). *Plant Dis.* 104, 1487–1491. doi: 10.1094/pdis-08-19-1602-re
- Sharma, R., Upadhyaya, H. D., Manjunatha, S. V., Rai, K. N., Gupta, S. K., and Thakur, R. P. (2013). Pathogenic variation in the pearl millet blast pathogen *Magnaporthe grisea* and identification of resistance to diverse pathotypes. *Plant Dis.* 97, 189–195. doi: 10.1094/pdis-05-12-0481-re
- Sillero, J. C., Rojas-Molina, M. M., Emeran, A. A., Kharrat, M., Winkler, J., Khan, H. R., et al. (2017). Identification and multi-environment validation of resistance to rust (*Uromyces viciae-fabae*) in *Vicia faba*. *Crop Pasture Sci.* 68, 1013–1023. doi: 10.1071/cp17099
- Silva, J., Groth, D. E., Moldenhauer, K. A., and Oard, J. H. (2011). GGE biplot exploration of resistance to sheath blight disease in doubled-haploid lines of rice. *Crop Sci.* 51, 1028–1035. doi: 10.2135/cropsci2010.10.0612
- Singh, B., Das, A., Parihar, A. K., Bhagawati, B., Singh, D., Pathak, K. N., et al. (2020). Delineation of genotype-by-environment interactions for identification and validation of resistant genotypes in mungbean to root-knot nematode (*Meloidogyne incognita*) using GGE biplot. *Sci. Rep.* 10:4108.
- Singh, S., Sharma, R., Pushpavathi, B., Gupta, S. K., Durgarani, C. V., and Raj, C. (2018). Inheritance and allelic relationship among gene (s) for blast resistance in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Plant Breed.* 137, 573–584. doi: 10.1111/pbr.12607
- Thakur, R. P., Sharma, R., Rai, K. N., Gupta, S. K., and Rao, V. P. (2009). Screening techniques and resistance sources for foliar blast in pearl millet. *J. SAT Agric. Res.* 7, 1–5.
- Thakur, R. P., Sharma, R., Rao, V. P., and Rajan, S. (2011). *Screening Techniques for Pearl Millet Diseases*. Patancheru: International Crops Research Institute for the Semi-Arid Tropics.
- Yan, W. (2001). GGE biplot – a windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agron. J.* 93, 1111–1118. doi: 10.2134/agronj2001.9351111x
- Yan, W. (2002). Singular–value partitioning in biplot analysis of multi-environment trial data. *Agron. J.* 94, 990–996. doi: 10.2134/agronj2002.0990
- Yan, W. (2011). GGE biplot vs. AMMI graphs for genotype-by-environment data analysis. *J. Ind. Soc. Agric. Stat.* 65, 181–193.
- Yan, W., and Falk, D. E. (2002). Biplot analysis of host-by-pathogen data. *Plant Dis.* 86, 1396–1401. doi: 10.1094/pdis.2002.86.12.1396
- Yan, W., and Holland, J. B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica* 171, 355–369. doi: 10.1007/s10681-009-0030-5
- Yan, W., Hunt, L. A., Sheng, Q., and Szlavics, Z. (2000). Cultivar evaluation and mega–environment investigation based on the GGE biplot. *Crop Sci.* 40, 597–605. doi: 10.2135/cropsci2000.403597x
- Yan, W., and Kang, M. S. (2003). *GGE biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists*. Boca Raton, FL: CRC Press.
- Yan, W., Kang, M. S., Ma, B., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype–by–environment data. *Crop Sci.* 47, 643–653. doi: 10.2135/cropsci2006.06.0374
- Yan, W., and Rajcan, I. (2002). Biplot evaluation of test locations and trait relations for breeding superior soybean cultivars in Ontario. *Crop Sci.* 42, 11–20. doi: 10.2135/cropsci2002.0011
- Yan, W., and Tinker, N. A. (2005). An integrated biplot analysis system for displaying, interpreting, and exploring genotype \times environment interaction. *Crop Sci.* 45, 1004–1016. doi: 10.2135/cropsci2004.0076
- Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: principles and applications. *Can. J. Plant Sci.* 86, 623–645. doi: 10.4141/p05-169

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sankar, Singh, Prakash, Satyavathi, Soumya, Yadav, Sharma, Rao, Singh and Srivastava. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Multi-Environment Quantitative Trait Loci Mapping for Grain Iron and Zinc Content Using Bi-parental Recombinant Inbred Line Mapping Population in Pearl Millet

Tripti Singhal¹, C. Tara Satyavathi^{2*}, S. P. Singh^{1*}, Aruna Kumar³, S. Mukesh Sankar¹, C. Bhardwaj¹, M. Mallik¹, Jayant Bhat⁴, N. Anuradha⁵ and Nirupma Singh¹

¹ ICAR-Indian Agricultural Research Institute, New Delhi, India, ² ICAR-All India Coordinated Research Project on Pearl Millet, Jodhpur, India, ³ Amity Institute of Biotechnology, Amity University, Noida, India, ⁴ Regional Research Centre, ICAR-Indian Agricultural Research Institute, Dharwad, India, ⁵ Acharya N. G. Ranga Agricultural University, Vizianagaram, India

OPEN ACCESS

Edited by:

Velu Govindan,
International Maize and Wheat
Improvement Center, Mexico

Reviewed by:

Sushil Kumar,
Anand Agricultural University, India
Mahesh Damodhar Mahendrakar,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

*Correspondence:

C. Tara Satyavathi
csatyavathi@gmail.com
S. P. Singh
sumerpalsingh@yahoo.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 28 January 2021

Accepted: 06 April 2021

Published: 18 May 2021

Citation:

Singhal T, Satyavathi CT,
Singh SP, Kumar A, Sankar SM,
Bhardwaj C, Mallik M, Bhat J,
Anuradha N and Singh N (2021)
Multi-Environment Quantitative Trait
Loci Mapping for Grain Iron and Zinc
Content Using Bi-parental
Recombinant Inbred Line Mapping
Population in Pearl Millet.
Front. Plant Sci. 12:659789.
doi: 10.3389/fpls.2021.659789

Pearl millet is a climate-resilient, nutritious crop with low input requirements that could provide economic returns in marginal agro-ecologies. In this study, we report quantitative trait loci (QTLs) for iron (Fe) and zinc (Zn) content from three distinct production environments. We generated a genetic linkage map using 210 F₆ recombinant inbred line (RIL) population derived from the (PPMI 683 × PPMI 627) cross using genome-wide simple sequence repeats (SSRs). The molecular linkage map (seven linkage groups) of 151 loci was 3,273.1 cM length (Kosambi). The content of grain Fe in the RIL population ranged between 36 and 114 mg/Kg, and that of Zn from 20 to 106 mg/Kg across the 3 years (2014–2016) at over the three locations (Delhi, Dharwad, and Jodhpur). QTL analysis revealed a total of 22 QTLs for grain Fe and Zn, of which 14 were for Fe and eight were for Zn on three consecutive years at all locations. The observed phenotypic variance (R^2) explained by different QTLs for grain Fe and Zn content ranged from 2.85 (QGFe.E3.2014–2016_Q3) to 19.66% (QGFe.E1.2014–2016_Q3) and from 2.93 (QGZn.E3.2014–2016_Q3) to 25.95% (QGZn.E1.2014–2016_Q1), respectively. Two constitutive expressing QTLs for both Fe and Zn co-mapped in this population, one on LG 2 and second one on LG 3. Inside the QTLs candidate genes such as Ferritin gene, Al³⁺ Transporter, K⁺ Transporters, Zn²⁺ transporters and Mg²⁺ transporters were identified using bioinformatics approaches. The identified QTLs and candidate genes could be useful in pearl millet population improvement programs, seed, restorer parents, and marker-assisted selection programs.

Keywords: pearl millet, QTL mapping, iron and zinc content, SSR, RILs

INTRODUCTION

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is a climate-resilient and health-promoting nutritious crop of the semi-arid tropics of Africa and Asia (Satyavathi et al., 2015). It is a staple food crop for approximately 90 million people that live in Asia and Africa and practice low-input subsistence farming and livestock production systems (Gangashetty et al., 2016). Pearl millet is an ideal crop

in the context of global climate change due to its inherent tolerance to heat stress, salinity, and drought (Shivhare and Lata, 2016). Deficiencies or imbalanced intakes of energy or nutrients, particularly vitamins and minerals, cause a number of dysfunctions and diseases in humans, that are collectively referred to as “micronutrient malnutrition or hidden hunger.” Currently more than 2 billion people in the world are in risk of micronutrient malnutrition (Athar et al., 2020). Iron (Fe) and Zinc (Zn) deficiency in humans is caused by lack of purchasing power for highly nutritious food, reduced dietary intake and less bioavailability and bioutilization, especially among resource poor women, infants, and children in the developing countries (Shariatipour and Heidari, 2020). Nutrition supplementation, dietary diversification, and food fortification are among the targeted strategies available to address micronutrient deficiencies. Among, the above strategies, improving the nutrient profile of agricultural crops through genetic means is more pleasing as it is a one-time investment and becomes self-sustaining over a long period. Breeding of pearl millet varieties/hybrids with improved nutritional quality is one of the priority areas for providing nutritional security in the developing countries. Many breeders (Govindaraj et al., 2011, 2016; Rai et al., 2015; Satyavathi et al., 2015; Kumar et al., 2016) have identified huge variability for the trait and utilizing them in development of a handful of varieties and hybrids rich in micronutrients, which now playing a major role in nutritional security in dry lands. However, hybrid released so far in India is having as mean Fe content at par or less than HarvestPlus target level of 77 mg kg^{-1} (HarvestPlus, 2014). Underutilization of germplasm material with higher nutritional value may attribute due to poor agronomic qualities. Hence there is a great need to make constant effort to breed more and more parental lines which can combine both for yield and micronutrient content along with resistance/tolerance to drought, downy mildew and blast diseases, good keeping quality, etc., and showing a stable performance under different agro-geographical regions.

Genetic maps provide important information for detailed genetic analysis of quantitative traits and have been proven to be significant method for crop improvement (Mohan et al., 1997; Doerge, 2002). Identifying molecular markers which are closely associated with the nutritional trait such as Fe and Zn enable rapid introgression of such traits into elite backgrounds. Among the different marker systems, simple sequence repeat (SSRs) is considered as a powerful and practically useful marker system for marker-assisted breeding (MAB) because of its cheap and more user-friendly along with co-dominant nature, abundance, reproducibility, and variability (Wang et al., 2018). In addition, the size and type of the mapping population determine the gene effect for an economically important trait (Kumar et al., 2018). Recombinant Inbred Lines (RILs) are the best choice among different types of mapping population as they are homozygous as well as have undergone multiple cycles of recombination which helps in mapping of tightly linked markers. Further use of codominant markers along with large RIL population helps in the construction accurate high-resolution map.

There are so many examples that SSR markers being employed for marker assisted biofortification of crop varieties

such as QPM and pro-vitamin rich maize hybrids (Gupta et al., 2009; Muthusamy et al., 2014) and development of high yielding basmati varieties with disease resistance (Singh et al., 2012).

There are two different strategies that the plant can able to uptake iron from the rhizosphere: Strategy I and Strategy II (Connorton et al., 2017). Most of the cereals like rice, maize, and wheat follow Strategy II through secretion of phytosiderophores (PSs) (Roberts et al., 2004). *TOM1* (Transporter of Mugineic acid family phytosiderophores)/*ZIFL4* is a member of major facilitator superfamily (MFS) which exports Fe^{3+} -PS chelates in rice and barley (Nozoye et al., 2011). In addition to Strategy II, rice and barley both have a functional Iron-Regulated Transporter 1 (IRT1) homolog that allows direct Fe^{2+} uptake from the rhizosphere. Metal transporters, such as the P1B-ATPase family, zinc-regulated transporter (ZRT), iron-regulated transporter (IRT)-like protein (ZIP), natural resistance-associated macrophage protein (NRAMP) family, and cation diffusion facilitator (CDF) family, have been reported in plants, where these two metals play critical roles (Colangelo and Guerinot, 2006). Several genes/gene families involved in Fe and Zn homeostasis such as yellow stripe-like (YSL), ZRT, ZIP, NRAMP, nicotianamine synthase (NAS), nicotianamine aminotransferase (NAAT), heavy metal ATPases (HMA), zinc-induced facilitator-like (ZIFL), metal tolerant protein (MTP), and others have been characterized in cereal crops such as rice (Anuradha et al., 2012), sorghum (Kotla et al., 2019). In pearl millet, *PgLYSL*, *PgZIP*, and *PgINRAMP* and a single gene for *PgIFER* and *PgINAS* were identified (Mahendrakar et al., 2020). Hence, the aim of this study was to identify the QTL associated with grain iron and zinc content through construct of linkage map using SSRs markers in a large-sized RIL mapping population.

MATERIALS AND METHODS

Plant Materials

The experimental materials comprised of 210 $F_{6:7}$ RILs mapping population derived from an intra specific cross between PPMI 683 (P_1 , Female) and PPMI 627 (P_2 , Male) made during rainy season, 2009. The RIL population advanced by single seed descent method from F_2 to F_6 from single plants selected during F_2 , and was segregating for grain Fe and Zn contents, respectively. Parent (P_1), i.e., PPMI 683 having high grain iron and zinc content and parent (P_2), i.e., PPMI 627 having low grain iron and zinc content with checks (ICMB 98222 for high grain Fe and Zn content, ICMB 92111 for low grain Fe and Zn content) (Table 1). Both checks were planted after every 20th row of the population.

The average grain yield of five randomly selected plants was taken as grain yield per plant and weight of randomly counted 1,000 seeds in each genotype was taken as 1,000 Seed Weight. This study is mainly focused on grain Fe and Zn content. Hence, only the two most important traits viz., seed yield per plant and 1,000 seed weight were considered for the study.

Environments

A total of 210 RILs along with both parents and checks were evaluated in an alpha lattice design with two replicates (Patterson and Williams, 1976), under field conditions at three different geographical areas, showing all three pearl millet growing agro-climatic zones of India, (i) E1: ICAR-Indian Agricultural Research Institute, New Delhi (28°38'N, 77°80'E) representing Zone A with annual rainfall of more than 400 mm, (ii) E2: ICAR-IARI Regional Centre farm, Dharwad (15°21'N, 75°05'E) from zone B (covering the southern peninsular India), and (iii) E3: Agricultural Research Station, All India coordinated pearl millet project, Mandor, Jodhpur (26°25'N, 72°99'E) falling in zone A₁ with annual rainfall <400 mm. The trials were conducted at all locations during three consecutive (south-west monsoon season, 2014, 2015, and 2016) growing seasons. Each RIL was sown on a four-rows plot of 4 m long with an inter-row spacing of 65 cm and plant-to-plant spacing of 10 cm. For a normal, healthy crop, standard pearl millet cultivation practices were strictly followed. **Table 2** shows the weather and soil parameters at each location throughout the year. All climatic parameters except rainfall are presented as means over the crop-growing season (June–October). Rainfall is cumulative rainfall received during the crop-growing season. The relative humidity was calculated as

the average of measurements taken in every morning and afternoon on each day.

Sample Preparation and Extraction of Mineral Elements

At physiological maturity, panicles from five plants per plot were randomly selected from the central two rows and harvested separately and threshed using wooden mallet. Seeds from individual plants within the plot were mixed and washed with deionized double-distilled water to remove debris and other possible contaminations in order to obtain a representative sample for trait measurements. The grain samples were sealed in individual paper bags and placed in an oven at 75°C overnight. About 15 g of air-dried sample from each genotype was further used. Mineral (Fe and Zn) content in the grains during initial year (2014) was measured with EDXRF instrument (a non-destructive, bench-top, energy-dispersive X-ray fluorescence spectrometry, Oxford Instruments X-Supreme 8000) at all three locations (Anuradha et al., 2017). Later, in 2015–2016, a more precise method, di-acid digestion of ground flour samples was used (Dhyan et al., 2005), followed by readings taken on an atomic absorption spectrometer (AAS, ZEE nit 700 tech Analytikjena). To determine whether higher grain mineral content is due to contamination with dust, soil, or is inherent,

TABLE 1 | Pedigree of the parental lines used to develop F_{6:7} mapping populations along with the high and low iron checks.

Parent	Iron (mg/kg)	Zinc (mg/kg)	Pedigree and developer
PPMI 683	103	72	High iron rich restorer line developed from in-iadi germplasm resistant for downy mildew at IARI, New Delhi
PPMI 627	49	28	Promising medium maturity restorer derived through selection in Pusa composite 334 at IARI, New Delhi
ICMB 98222 (high Fe check)	85	45.8	ARD-288-1-10-1-2(RM)-5, developed by ICRISAT, India
ICMB 92111 (low Fe Check)	43.6	27.5	(81B × 843B)-11-1-1-B, developed by ICRISAT, India

TABLE 2 | Geographical location, climatic, and edaphic factors over different years and locations.

Environment parameter	New Delhi (E1)			Dharwad (E2)			Jodhpur (E3)		
	2014	2015	2016	2014	2015	2016	2014	2015	2016
Geographical identity									
Latitude	28.70°N	28.70°N	28.70°N	15.45°N	15.45°N	15.45°N	26.35°N	26.35°N	26.35°N
Longitude	77.10°E	77.10°E	77.10°E	75.00°E	75.00°E	75.00°E	73.04°E	73.04°E	73.04°E
Altitude (m.s.l.)	219 m	219 m	219 m	751 m	751 m	751 m	231 m	231 m	231 m
Climatic									
Temp (max.)	34.98°C	34.48°C	33.88°C	27.8°C	29.13°C	26.77°C	34.4°C	36.8°C	33.08°C
Temp (min.)	24.08°C	23.58°C	22.38°C	20.47°C	20.76°C	20.47°C	25°C	25.6°C	24.76°C
Relative humidity (%)	70.2	70.7	75.2	80.9	78.06	82.81	63.5	60.3	62.28
Rainfall (mm)	472.6	717.3	1146.7	544.9	293.2	385.4	389	288	427
Soil factors									
Soil pH	7.89	8.1	7.87	7.1	6.8	6.9	8.4	8.2	8.2
Electrical conductivity (dSm ⁻¹)	0.23	0.25	0.22	0.18	0.2	0.2	0.09	0.1	0.08
% Organic content	0.26	0.24	0.25	0.31	0.29	0.3	0.22	0.21	0.27
Av. Fe (mg kg ⁻¹)	18.14	21.04	17.89	27.6	30.7	28.1	6.4	4.9	2.27
Av. Zn (mg kg ⁻¹)	2.3	4.62	2.12	1.7	2	1.9	3.2	2.9	1.35
Soil texture	Sandy loam	Sandy loam	Sandy loam	Silty clay	Silty clay	Silty clay	Loamy sand	Loamy sand	Loamy sand

ICPMS (Nex ION 300X, Perkin Elmer, United States) was used to analyze genotypes having high iron content ($>90 \text{ mg kg}^{-1}$) (Pfeiffer and McClafferty, 2007).

Statistical Analysis

Descriptive statistics and Analysis of Variance (ANOVA) was carried out using PBTools v1.4. Histograms and correlations between pairs of traits were estimated through Pearson correlation co-efficient using R software. The model used for ANOVA was:

$$z_{ijklm} = \mu + y_i + e_j + y_{ej} + r_{ijk} + b_{ijkl} + g_m + (yg)_{im} + (eg)_{jm} + (yeg)_{ijm} + \epsilon_{ijklm}$$

Where μ is the grand mean; y_i is the fixed effect of year i ; e_j is the fixed effect of location j ; y_{ej} is the fixed effect of interaction between year i and location j ; g_m is the random effect of genotype m and is $\sim \text{NID}(0, \sigma_g^2)$; r_{ijk} is the random effect of replication in location j and year i and is $\sim \text{NID}(0, \sigma_r^2)$; b_{ijkl} is the random effect of block l nested with replication k in location j and year i and is $\sim \text{NID}(0, \sigma_b^2)$; $(yg)_{im}$ is the random effect of the interaction between genotype m and year i and is $\sim \text{NID}(0, \sigma_{yg}^2)$; $(eg)_{jm}$ is the random effect of the interaction between accession m in location j and $\sim \text{NID}(0, \sigma_{eg}^2)$; $(yeg)_{ijm}$ is the random effect of the interaction effect of the genotype m in year i and location j and $\sim \text{NID}(0, \sigma_{yeg}^2)$; and ϵ_{ijklm} is the random residual effect and $\sim \text{NID}(0, \sigma_\epsilon^2)$.

Analysis of variance was also conducted using data from each environment for grain Fe and Zn content.

Heritability (H^2 , broad sense) at individual environment was estimated from analysis of variance. The formula used was-

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_\epsilon^2}{r}}$$

Whereas Heritability (H^2) estimates across environments were estimated by the formula-

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{yg}^2}{y} + \frac{\sigma_{eg}^2}{l} + \frac{\sigma_{yeg}^2}{yl} + \frac{\sigma_\epsilon^2}{ryl}}$$

Where r , y , l denotes the number of replicates, years and environments, respectively.

The REML model also produced best linear unbiased predictors (BLUPs) of each genotype, thereby adjusting the influence of the neighboring rows. These BLUPs were used for downstream analysis.

Genomic DNA Extraction and Genotyping

The genomic DNA was extracted from fresh and young leaf tissue of each RIL and the parents by using modified cetyltrimethylammonium bromide (CTAB) method (Aboul-Maaty and Oraby, 2019). Polymorphism was assessed using 372 SSR markers, EST-SSRs (ICMP and IPES series), gSSRs (PSMP and CTM series), and gene-based markers. SSRs were amplified in 10 μl PCR

reaction mixture containing 25–30 ng genomic DNA, 100 μM of each of the four dNTPs, 1.5 mM MgCl_2 , 5 pmol/ μl of forward and reverse primer, 1 U of Taq DNA polymerase (Bangalore-genei, Merck) and $1\times$ PCR buffer in Biorad My Cycler Thermal Cycler, United States. PCR conditions were used as described by Anuradha et al. (2017). The amplified product was checked using 1.2% agarose gel and PCR products were separated by 10% polyacrylamide gel electrophoresis at 120 V (C.B.S. Scientific vertical electrophoresis unit). The SSR marker alleles were scored manually by comparing the position of bands with that of 100 bp ladder from bottom to top (smallest to largest). Based on the position, the allele size in base pairs (bp) was recorded for all polymorphic SSRs in all 210 RILs studied. The parent P₁ (PPMI 683) was scored as “A” and parent P₂ (PPMI 627) was scored as “B” and both type band present on single locus were designed as “H”. Diffused or ambiguous bands were recorded as missing bands with “0” notation. The chi-square method was conducted to test whether the inherited alleles of the mapping population were consistent with Mendelian segregation ratios. The segregation ratio across the mapping population was tested against a 1:1 ratio. The segregation pattern that did not fit either ratio ($p < 0.05$) were treated as distorted.

Linkage Analysis and QTL Mapping

The linkage map was constructed using 151 polymorphic SSR markers. The markers that were not taken for the linkage analysis were those with: a minor allele frequency of less than 5%, more than 20% missing data, and those that were monomorphic between the parents. The MAPMAKER/EXP3.0 software was used to develop a genetic linkage map (Lander et al., 1987). The maximum recombination fraction (θ) was set at 0.49 and the critical logarithm of odds (LOD) score for the test of marker pair independence was set at 3. Markers with high segregation distortion ($p \leq 0.01$) were removed in a preliminary analysis of the data. Subsequently, the linkage analysis of genotyping data for all the markers including those with distorted segregation was used. Using MapChart software, a graphical map was generated (Voorrips, 2002). The consensus map developed by Supriya et al. (2011) and Rajaram et al. (2013) was used as a reference map in the current study for grouping markers and constructing the map. QTL analysis was carried out using the Inclusive Composite Interval Mapping method (ICIM). The ICIM approach uses a strategy in which a stepwise regression is firstly performed, so markers with significant effect on QTL are selected. ICIM-ADD method was used with the multi-environmental model builtin QTL IciMapping (Li et al., 2015). Significant LOD thresholds were set at 5% tail of the null distribution in a 1,000 permutations test (Luciano Da Costa et al., 2012).

Quantitative Trait Loci Validation

All QTL intervals identified for grain Fe and Zn content were analyzed *in-silico* with reference genome, Tift23D2B1-P1-P5 (PROJECT ID: PRJNA294988.) in order to navigate the presence of candidate gene associated with Fe and Zn content in grain. The physical position of flanked markers of identified QTLs

were obtained from NCBI-BLASTn¹. The sequence similarities searched against related cereals (*Setaria italica*) by Ab initio gene prediction method by using Fgenesh software for the identification of potential protein coding regions (Salamov and Solovyev, 2000)². For each predicted exon similarity searches were run using the BLAST program on protein and EST databases (Altschul et al., 1997).

RESULTS

Phenotypic Evaluations

The mean performance of parents and range of trait values of RILs during three consecutive years (2014, 2015, and 2016) at E1, E2, and E3 are summarized in **Table 3**. Both traits (Fe and Zn) were approximately normally distributed when analyzed across the years at their respective locations (**Figure 1**). The levels of grain Fe in RILs ranged from 43.27 to 108.33 mg kg⁻¹, 33.34–118.52 and 32.88–115.47 mg kg⁻¹, while the range of grain Zn content was 22.4–94, 20.97–120.87, and 17.33–105.17 mg kg⁻¹ across the years at E1, E2, and E3, respectively. The mean BLUP value of grain Fe and Zn content in both parents and RILs were higher across the described years at E1 when compared to E2 and E3, whereas mean BLUPs of grain Fe and Zn for both parents was lowest at E3 across the years. For grain Fe and Zn studied in different environments, both parents and RILs showed a wider range of variation.

The seed yield per plant and 1,000 seed weight of the 210 RILs along with both parents varied greatly. At E1 during three consecutive years (2014–2016), 1,000 seed weight ranged from 6.75 to 12.17 g with a mean of 9.05 g, and seed yield per plant ranged from 14.09 to 36.21 g with the mean of 24.05 g. And at E2 from 2014 to 2016, 1,000 seed weight ranged from 6.77 to 12.33 g with a mean of 9.09 g, and seed yield per plant ranged from 19.05 to 35.37 g with a mean of 26.81 g. From 2014 to 2016, 1,000 seed weights in E3 ranged from 6.48 to 11.41 g, with an average of 8.38 g, and seed yield per plant was 18.20–39.8 g, with an average of 26.75 g.

The ANOVA indicated a highly significant $G \times Y$ interaction [$p < 0.0001$, $F_{(66,45)} = 1.112$ for Fe] and [$p < 0.0001$, $F_{(62,43)} = 1.112$ for Zn]. The pooled mean, range, coefficient of variation (CV), genotypic variance (σ_g^2), residual variance, broad sense heritability (H^2), least square difference (LSD) and $G \times E$ variance in RILs evaluated during 2014–2016 cropping season at E1, E2, and E3 has shown in **Table 4**. In the RIL population, the genotypic components of variance (σ_g^2) for both traits were significant, and the operational heritability (H^2) estimates ranged from 0.83 (Zn) to 0.88 (Fe) (**Table 4**). The high heritability shows that much of the phenotypic variance in the population is genetically controlled and QTL can be mapped with a high degree of reliability. Broad-sense heritability was usually high enough to allow effective QTL mapping, indicating that the traits studied have moderately high to high proportions of genetic

TABLE 3 | Grain Fe and Zn content (mg Kg⁻¹) \pm SE of the parents along with RILs of Fe₆₇ mapping populations over 3 years and three diverse locations.

Locations		Delhi (E1)						Dharwad (E2)						Jodhpur (E3)							
		Iron			Zinc			Iron			Zinc			Iron			Zinc				
micron	nutrient	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year
Years		14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year	14-15	15-16	16-17	Across year
PPMI 683		98.5 ± 2.5	100.5 ± 2.5	98 ± 6	99 ± 0.76	83 ± 15	78.5 ± 3.5	91 ± 5	84.17 ± 3.66	95.56 ± 1.17	94.06 ± 1.33	96.56 ± 1.33	95.39 ± 0.73	70.37 ± 2	76.04 ± 2.67	83.54 ± 0.83	69.98 ± 3.61	83.45 ± 2.33	58.6 ± 2.67	59.77 ± 2	60.88 ± 1.73
PPMI 627		51 ± 1	53 ± 1	47 ± 1	50.33 ± 1.76	32 ± 0	33 ± 3	26.5 ± 0.5	30.5 ± 2.02	37.74 ± 1.67	35.07 ± 1.5	41.07 ± 2.17	37.96 ± 1.74	23.65 ± 2.5	21.15 ± 1.33	27.65 ± 1.83	24.15 ± 1.89	35 ± 0.8	34.33 ± 1.83	34.22 ± 2	18.5 ± 2.33
Pop (MEAN)		71.15 ± 1.01	71.45 ± 0.96	71.55 ± 0.96	71.38 ± 0.96	51.41 ± 1.01	55.31 ± 0.88	53.47 ± 0.93	53.39 ± 0.84	70.99 ± 1.18	70.92 ± 1.21	70.82 ± 1.21	70.91 ± 1.19	51.43 ± 0.89	49.67 ± 1.05	51.58 ± 1.02	50.9 ± 0.93	69.69 ± 1.03	69.78 ± 1.01	69.59 ± 1.05	69.69 ± 1.02

¹<https://blast.ncbi.nlm.nih.gov/>

²<http://www.softberry.com/berry.phtml?topic=fgenesh&group=programs&subgroup=gfind>

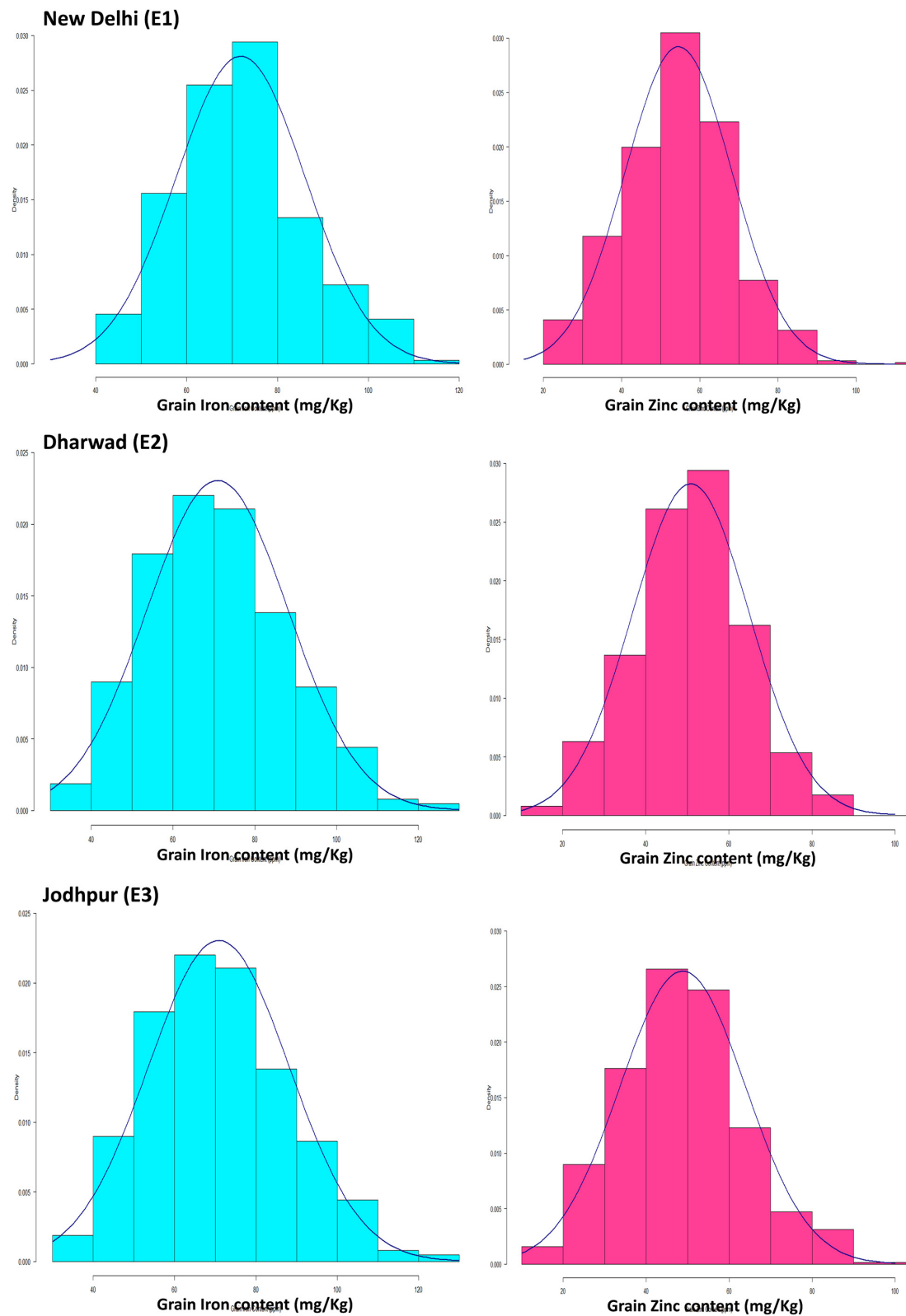


FIGURE 1 | Histograms of grain Fe and Zn in mapping populations of recombinant inbred lines evaluated across the year (2014–2016) at Delhi, Dharwad, and Jodhpur location.

TABLE 4 | Estimates of mean, variance components (2014, 2015, and 2016), broad-sense heritability and LSD at Delhi, Dharwad, and Jodhpur for grain Fe and Zn content in RILs.

Traits	Environment	Mean	Range (mg kg ⁻¹)	CV %	Genotype variance	Residual variance	Heritability	LSD	G × E variance
Fe									
1	New Delhi_2014	71.03 ± 1.00	41.4–111.5	4.47	211.25	10.08	0.88	4.66	70.00**
2	New Delhi_2015	72.71 ± 0.95	42.1–111	2.23	198.91	2.62	0.89	2.3	
3	New Delhi_2016	71.95 ± 0.96	44–108.5	2.44	203.28	3.09	0.89	2.69	
4	New Delhi across years	71.89 ± 0.96	43.27–108.33	3.19	161.09	5.27	0.81	7.51	
5	Dharwad_2014	70.95 ± 1.18	34.63–119.35	3.79	293.4	7.23	0.89	3.75	
6	Dharwad_2015	70.86 ± 1.21	17.71–125.7	3.99	311.4	8.01	0.89	3.94	
7	Dharwad_2016	70.81 ± 1.21	33.8–123.19	4.17	308.43	8.7	0.89	4.11	
8	Dharwad across years	70.87 ± 1.19	33.34–119.57	3.96	301.49	7.86	0.89	3.62	
9	Jodhpur_2014	69.6 ± 1.03	32.6–114.8	4	224.01	7.74	0.88	4.01	
10	Jodhpur_2015	69.68 ± 1.01	36.47–118	4.26	214.31	8.82	0.88	4.12	
11	Jodhpur_2016	69.48 ± 1.05	15.8–100	4.15	231.69	8.3	0.88	4	
12	Jodhpur across years	69.58 ± 1.02	32.88–115.47	4.02	223.56	7.84	0.89	2.94	
Zn									
13	New Delhi_2014	50.85 ± 1.01	20.1–114	5.01	200.13	6.48	0.88	3.88	107.00**
14	New Delhi_2015	58.79 ± 0.88	22.4–90	3.28	147.25	3.72	0.89	2.73	
15	New Delhi_2016	54.1 ± 0.94	21.2–89	3.18	188.94	2.96	0.89	2.64	
16	New Delhi across years	54.58 ± 0.84	22.4–94	3.84	110.81	4.38	0.83	8.72	
17	Dharwad_2014	51.39 ± 0.88	20.97–114.2	5.2	164.6	7.13	0.88	3.78	
18	Dharwad_2015	49.66 ± 1.04	17.71–125.7	5.71	230.7	8.04	0.88	3.94	
19	Dharwad_2016	51.52 ± 1.01	18.99–122.7	5.49	213.51	8	0.88	4	
20	Dharwad across years	50.86 ± 0.93	20.97–120.87	5.46	176.41	7.7	0.85	6.3	
21	Jodhpur_2014	48.52 ± 1.04	17.6–102	5.68	227.84	7.58	0.88	3.84	
22	Jodhpur_2015	49.72 ± 1.01	15.8–100	5.79	216.96	8.29	0.88	4	
23	Jodhpur_2016	48.19 ± 1.09	15–113.5	5.77	253.4	7.73	0.88	3.87	
24	Jodhpur across years	48.81 ± 1.03	17.33–105.17	5.74	224.59	7.84	0.88	4.32	

Significance at ** $p < 0.01$.

variance. Correlation coefficients (r) for grain Fe and Zn content were found statistically significant ($p < 0.001$) for each year of evaluations for each location.

Homogeneity of variance tests based on Bartlett's test indicated homogeneous error variance for both the traits which allowed for a combined analysis across the years for all locations. When performing a combined ANOVA, the year was considered as a random and genotypes were considered as fixed effects. Combined ANOVA showed significant Genotype × Environment Interaction (GEI) ($p < 0.001$), exhibiting the influence of changes in environment on the grain micronutrient contents of genotypes. It was observed that Genotypes (G) and GEI effects were highly significant ($p < 0.01$) and the contribution of year is very less for both the traits (Table 5).

Correlation of Grain Fe, Zn, 1,000 Seed Weight and Yield

Grain Fe and Zn content were highly (0.711) and significantly ($p < 0.01$) correlated (Table 6). Thousand seed weight was also observed to be positively correlated with seed yield per plant

(0.579, $p < 0.01$). For successful incorporation of high levels of both grain Fe and Zn content, association between these micronutrients and their association with thousand seed weight and yield is important to plan a successful breeding approach so as to develop new improved lines or varieties having higher grain yield along with enhanced levels of grain micronutrients.

Linkage Map Construction

For the polymorphism survey of both parental lines, 366 SSR primer pairs (IPES, PSMP, and ICMP series) and six gene-based primers were used. Of them, 151 (40.59%) SSRs detected polymorphism between the two parents. Polymorphic SSR markers included 94 IPES, 42 PSMP, and 11 ICMP series primer pairs and four newly synthesized primer pair named as Ppmsb (Pusa pearl millet SSR biofortification) (Figure 2). These polymorphic primers were further used for linkage map construction with LOD score 3.0.

The total length of the linkage map was 3273.1cM (Kosambi). The individual linkage group ranged from 910 cM for LG 2 with the highest number of markers (31) to 135.3 cM for LG 3 with 14 markers, while the least number of markers were

TABLE 5 | Mean sum of square of Fe and Zn at all locations based on combined analysis over three years for RILs and parents.

Sources	Degrees of freedom	Delhi_Fe	Delhi_Zn	Dharwad_Fe	Dharwad_Zn	Jodhpur_Fe	Jodhpur_Zn
Genotypes	211	1,059.8**	806**	1,823**	1,120**	1,349.8**	1,372**
Environments (years)	2	300.7	6,755**	2	456	4.4	274
Block	3	2,175.9**	1,765**	7,044**	6,747**	6,652**	7,290**
Year × genotype	422	92.9**	141**	14**	61**	7.9	24**
IPCA 1	212	108.5**	148**	26**	108**	13.1**	37**
IPCA 2	210	77.1**	134**	2	14**	2.7	11**
Total	1271	215.3	198	328	227	246.6	257

IPCA, interactive principal component analysis.

Significance at ** $p < 0.01$.

assigned on LG4 (10) followed by LG 3 (14) and LG 5 (14). The average interval size was 8.79 cM. The average LG length was 467.5 cM with an average of 21.57 loci. The average adjacent-marker interval lengths ranged from 3.07 cM (LG 3) to 14.91 cM (LG 2) followed by 11.34 cM (LG 1). Map distance between adjacent markers varied from 0.9 to 49.2 cM (LG 1) and 8.33% of the intervals (31 out of 372 intervals) were smaller than 10 cM. Details of all linkage group, total markers, skewed marker and their percentage, total map length (cM) and average distance (cM) were given in **Table 7**.

Quantitative Trait Loci Analysis

The multi-environmental QTL analysis exhibited the association of various genomic regions with grain Fe and Zn content at all the three locations during three consecutive years, 2014, 2015, and 2016 (**Table 8**). The analysis indicated the presence of four significant common QTL for grain Fe content, flanked by following markers Ppmsb001–Ppmsb002, IPES0142–IPES0180, IPES0157–IPES0093, and IPES0206–IPES0015, which explained the total phenotypic variance of 18.34, 17.98, and 15.98% for QTL1 on LG 2, 14.99, 16.26, and 14.8% for QTL 2 on LG 3, 19.66, 18.83, and 16.85% for QTL 3 on LG 5, 15.66, 16.61, and 14.56% for QTL 4 on LG 7 across three consecutive years (2014–2016) at E1, E2, and E3, respectively. These four QTLs were named as: *QGF_{E1}.2014–2016_Q1*, *QGF_{E1}.2014–2016_Q2*, *QGF_{E1}.2014–2016_Q3*, and *QGF_{E1}.2014–2016_Q4* for Delhi location (E1), *QGF_{E2}.2014–2016_Q1*, *QGF_{E2}.2014–2016_Q2*, *QGF_{E2}.2014–2016_Q3*, and *QGF_{E2}.2014–2016_Q4* for Dharwad location (E2), and *QGF_{E3}.2014–2016_Q1*, *QGF_{E3}.2014–2016_Q2*, *QGF_{E3}.2014–2016_Q3*, and

QGF_{E3}.2014–2016_Q4 for Jodhpur location (E3) (Q = QTL; G = grain; Fe = iron). In addition, two more QTLs for GFe were also found at E3, flanked by IPES0101–IPES0139 and IPES0027–IPES0236 and having very less PVE of 2.92 and 2.85% on LG 1 and LG 2, respectively. The genome wide significant threshold was LOD = 3.5 for E1, E2, and E3.

For GZn, one common QTL was found in E1, E2, and E3 on LG 3 at significant threshold LOD value of 3.5 which was flanked by IPES0142–IPES0180 with total phenotypic variance of 14.39, 16.07, and 22.26% at E1, E2, and E3, respectively. And one more common QTL was also found at E1 and E2 on LG 2, flanked by Ppmsb001–Ppmsb002 having total PVE of 25.95 and 21.88%, respectively. Whereas, at E3, one QTL with flanked markers Ppmsb002–IPES0181 was mapped on LG 2 with comparatively low total PVE of 11.18%. Two more un-repetitive QTLs between IPES0166–PSMP2249 (PVE 6.2%) on LG 3 and IPES0141–IPES0151 (PVE 2.93%) on LG 6 were detected in E1 and E3, respectively.

The proportion of the total phenotypic variance (PVE) explained by the QTLs for GFe ranged from 14.99 to 19.66%, 16.26–18.83, and 2.85–16.85% in E1, E2, and E3, respectively (**Table 8**). On the other hand, the PVE of the QTLs for GZn in E1, E2, and E3 ranged from 6.2 to 25.95%; 16.07–21.88 and 2.93–22.26% (**Table 8**). The QTLs that had the highest PVE for GFe in E1, E2, and E3 were 19.66% (IPES0157–IPES0093), 18.83% (IPES0157–IPES0093), and 16.85% (IPES0157–IPES0093), respectively. For GZn, the QTLs with the highest PVE were 25.95% (Ppmsb001–Ppmsb002), 21.88% (Ppmsb001–Ppmsb002), and 22.26% (IPES0142–IPES0180) at E1, E2, and E3, respectively.

We localized two genomic regions with pleiotropic effects on LG 2 and LG 3 for GFe and GZn at E1, E2, and E3. The first genomic region was *QGF_{E1}.2014–2016_Q1* to *QGZn.E1.2014–2016_Q1* at E1, *QGF_{E2}.2014–2016_Q1* to *QGZn.E2.2014–2016_Q1* at E2 and *QGF_{E3}.2014–2016_Q2* to *QGZn.E3.2014–2016_Q1* at E3 linked to marker Ppmsb1–Ppmsb2 on LG 2 (**Table 8**). The second region was associated with *QGF_{E1}.2014–2016_Q2* to *QGZn.E1.2014–2016_Q3*, *QGF_{E2}.2014–2016_Q2* to *QGZn.E2.2014–2016_Q2*, and *QGF_{E3}.2014–2016_Q4* to *QGZn.E3.2014–2016_Q2* linked to marker IPES0142–IPES0180, across the years, at all the three locations.

The *QGF_{E1}.2014–2016_Q3* region linked to marker IPES0157–IPES0093 having largest PVE for GFe (19.66%) on

TABLE 6 | Correlation of grain iron and zinc content with thousand seed weight and grain yield.

	Fe	SYPP	TSW	Zn
Fe	1	0.049	0.032	0.711**
SYPP		1	0.579**	0.059
TSW			1	0.059
Zn				1

Fe, SY, TSW, and Zn are Grain iron content in mg/kg, Seed yield per plant in g, thousand seed weight in g, and Grain zinc content in mg/kg.

**Significance at $p < 0.01$.

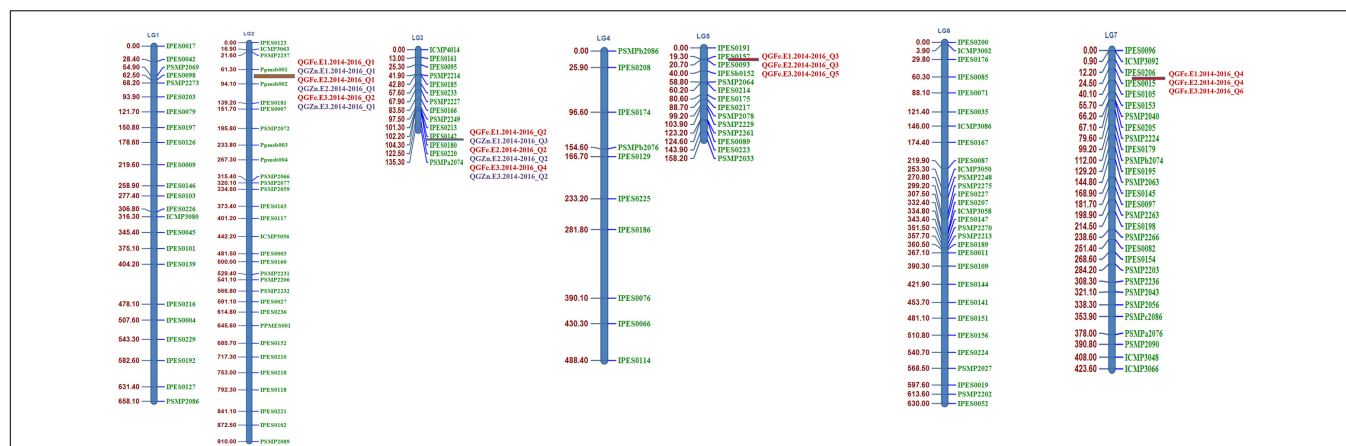


FIGURE 2 | Linkage map (linkage groups LG 1–LG 7) of the pearl millet recombinant inbred line (RIL) population based on the cross (PPMI 683 × PPMI 627). QTL identified in red color font indicates the grain Fe QTLs and QTL identified in blue color font indicates the grain Zn QTLs.

TABLE 7 | Details of the simple sequence repeats (SSRs)-based linkage map of the pearl millet RIL population based on the cross PPMI 683 × PPMI 627.

Linkage group	SSR marker loci	Skewed SSRs	Total marker loci	Skewed loci %	Total length (cM)	Average distance (cM)
1	23	35	58	60.34	658.1	11.34
2	31	30	61	49.18	910	14.91
3	14	30	44	68.18	135.3	3.075
4	10	30	40	75	361.8	9.045
5	14	35	49	71.42	158.2	3.22
6	30	30	60	50	626.1	10.43
7	29	31	60	51.6	423.6	7.06
Total	151	221	372	425.72	3,273.1	59.08
Average per linkage group	21.57	31.57	53.14	59.4	467.5	8.79

LG 5 over the locations, which was also displayed 3.6 additive effect of QTL (Table 8). Whereas, the QTL for GZn having largest PVE (25.95%) was *QGZn.E1.2014–2016_Q1* on LG 2 over the locations (Table 8). Similarly, one more QTL for grain Zn (*QGZn.E2.2014–2016_Q1*) was also found with high PVE 21.88% on LG 2 at E2. The additive effect was detected for all the QTLs on which the alleles from PPMI 683 (donor parent) had positive additive effect.

Validation of Identified QTL in Diverse Set of Germplasm

Eleven high (PPMI 708, PPMI 1102, PPMFeZMP 199, PPMI 1278, 841B, PPMI 1232, PPMI 1107, PPMI 1225, PPMI 1089, PPMWGI 146, and PPMI 1231), nine low (PPMI 1011, PPMI 759, J 2467, PPMI 1087, PPMFeZMP 47, PPMFeZMP 34, J2405, and H77/833-2, 5540B) and four checks (ICMB 98222, ICMB 93222, ICMB 07999, and ICMB 92111) for grain iron expressing pearl millet genotypes (Table 9) were selected and their DNA was used for amplification with ten consistently associated primers, IPES0142, IPES0180, IPES 0160, IPES0093, IPES0206, IPES0015, IPES0166, PSMP2249, Ppmsb001, and Ppmsb002. The amplicons of these genotypes were obtained by PCR amplification of these primers. The similar regions of reference sequence obtained were thus compared with the amplified sequences of high and low

grain iron of expressing pearl millet genotypes. These markers were compared with reference genome Tift₂₃D₂B₁.

In-silico Identification of Candidate Gene

Candidate genes identified for each QTL was shown in Table 10. In the study, we identified ten metal transporter genes. Out of the ten genes seven were identified in the genome of *Seteria italica* whereas remaining three genes were identified in the genome of *Zea mays*. Among the genes from the *Seteria italica*, Ferritin-1 was associated with the QTL *QFe 2.1* and *QZn 2.1*, Aluminum-activated malate transporter 5 and Potassium transporter 25 was identified with *QFe 3.1* and *QZn 3.1*, Probable magnesium transporter NIPA4, Metal transporter Nramp3 isoform X2, Probable cadmium/zinc-transporting ATPase HMA1, chloroplastic isoform X2 and Cadmium /zinc-transporting ATPase HMA2 were linked with *QFe 5.1*. However, among the genes from the *Zea mays*, Potassium transporter 3 was associated with *QFe 5.1* and *QFe 7.1*.

DISCUSSION

Variance Components and Heritability

Pearl millet has a great variation of mineral and agronomic traits. For both traits, the parents of this RIL population exhibited

TABLE 8 | Quantitative trait loci for grain Fe (GFe) and Zn (GZn) content in RIL over three locations.

QTL name	LG	Position (cM)	Left marker	Right marker	Intermarker distance (cM)	LOD	LOD	LOD	PVE	PVE	PVE	Add	AbyE_01	AbyE_02	AbyE_03	Left CI	RightCI
							(A)	(A by E)		(A)	(A by E)						
<i>QGFe.E1.2014–2016_Q1 (QFe2.1)</i>	2	138.9	Ppmsb 001	Ppmsb 002	32.8	23.16	23.15	0	18.34	18.27	0.07	3.46	0.26	−0.26	0	135.9	141.9
<i>QGFe.E1.2014–2016_Q2 (QFe3.1)</i>	3	121	IPES0142	IPES 0180	2.1	19.11	19.1	0.01	14.99	14.97	0.01	3.13	0.08	−0.13	0.06	110	124
<i>QGFe.E1.2014–2016_Q3 (QFe5.1)</i>	5	21.3	IPES 0157	IPES 0093	1.4	24.92	24.69	0.23	19.66	19.58	0.09	3.58	−0.16	0.34	−0.18	19.3	30.3
<i>QGFe.E1.2014–2016_Q4 (QFe 7.1)</i>	7	24.9	IPES 0206	IPES 0015	12.3	20.06	20.03	0.02	15.66	15.61	0.05	3.21	0.25	−0.1	−0.14	21.9	27.9
<i>QGZn.E1.2014–2016_Q1 (QZn 2.1)</i>	2	138.9	Ppmsb 001	Ppmsb 002	32.8	21.39	20.85	0.54	25.95	24.72	1.23	4.51	1.42	−0.83	−0.58	133.9	141.9
<i>QGZn.E1.2014–2016_Q2 (QZn 3.1)</i>	3	101	IPES 0166	Xpsmp 2249	14.0	5.86	1.74	4.11	6.2	1.92	4.28	1.26	−1.5	2.65	−1.15	98	102
<i>QGZn.E1.2014–2016_Q3 (QZn 3.2)</i>	3	121	IPES 0142	IPES 0180	2.1	11.97	8.44	3.53	14.39	9.53	4.86	2.8	1.36	−2.83	1.47	108	124
<i>QGFe.E2.2014–2016_Q1(QFe2.1)</i>	2	138.9	Ppmsb 001	Ppmsb 002	32.8	23.61	23.57	0.05	17.98	17.95	0.02	4.22	−0.09	0.2	−0.11	133.9	141.9
<i>QGFe.E2.2014–2016_Q2 (QFe3.1)</i>	3	121	IPES 0142	IPES 0180	2.1	21.32	21.26	0.06	16.26	16.25	0.01	4.02	−0.03	0.13	−0.1	114	124
<i>QGFe.E2.2014–2016_Q3 (QFe5.1)</i>	5	21.3	IPES 0157	IPES 0093	1.4	24.8	24.61	0.19	18.83	18.77	0.06	4.32	0.06	−0.32	0.26	19.3	32.3
<i>QGFe.E2.2014–2016_Q4 (QFe 7.1)</i>	7	24.9	IPES 0206	IPES 0015	12.3	22.02	21.97	0.05	16.61	16.61	0	4.08	−0.08	0.03	0.06	21.9	27.9
<i>QGZn.E2.2014–2016_Q1(QZn 2.1)</i>	2	138.9	Ppmsb 001	Ppmsb 002	32.8	24.89	24.58	0.31	21.88	21.77	0.11	5.05	−0.46	0.39	0.07	135.9	139.9
<i>QGZn.E2.2014–2016_Q2 (QZn 3.1)</i>	3	121	IPES 0142	IPES 0180	2.1	18.86	18.39	0.46	16.07	16.05	0.02	4.34	−0.07	0.22	−0.14	106	124
<i>QGFe.E3.2014–2016_Q1(QFe 1.1)</i>	1	452.4	IPES 0101	IPES 0139	29.1	3.94	3.93	0.01	2.92	2.91	0	−1.56	−0.06	0.09	−0.03	435.4	461.4
<i>QGFe.E3.2014–2016_Q2 (QFe2.1)</i>	2	138.9	Ppmsb 001	Ppmsb 002	32.8	20.63	20.61	0.01	15.98	15.98	0.01	3.54	0	−0.08	0.08	133.9	141.9
<i>QGFe.E3.2014–2016_Q3 (QFe2.2)</i>	2	644.9	IPES 0027	IPES 0236	23.7	3.93	3.92	0.01	2.85	2.85	0.01	−1.55	−0.06	0.09	−0.03	633.9	657.9
<i>QGFe.E3.2014–2016_Q4 (QFe 3.1)</i>	3	105	IPES 0142	IPES 0180	2.1	19.26	19.23	0.03	14.8	14.8	0	3.41	−0.02	0.01	0.01	104	112
<i>QGFe.E3.2014–2016_Q5 (QFe 5.1)</i>	5	21.3	IPES 0157	IPES 0093	1.4	21.76	21.73	0.02	16.85	16.85	0	3.64	−0.01	−0.03	0.03	19.3	34.3
<i>QGFe.E3.2014–2016_Q6 (QFe 7.1)</i>	7	24.9	IPES 0206	IPES 0015	12.3	18.97	18.96	0.01	14.56	14.55	0.01	3.39	0.08	−0.13	0.05	21.9	27.9
<i>QGZn.E3.2014–2016_Q1 (QZn 2.1)</i>	2	140.9	Ppmsb 002	IPES 0181	45.1	11.94	10.95	0.99	11.18	10.23	0.95	3.98	0.68	−1.7	1.01	139.9	143.9
<i>QGZn.E3.2014–2016_Q2 (QZn 3.1)</i>	3	105	IPES 0142	IPES 0180	2.1	25.42	24.39	1.02	22.26	21.08	1.19	5.71	−1	1.92	−0.92	104	122
<i>QGZn.E3.2014–2016_Q3 (QZn 6.1)</i>	6	478	IPES 0141	IPES 0151	27.4	3.71	3.55	0.16	2.93	2.87	0.06	−2.19	−0.32	−0.12	0.45	471	485

LOD, overall LOD score of the QTL; LOD (A), score of the main additive effect; LOD (AbyE), score of the Genotype \times Year interaction; PVE, overall proportion of phenotypic variance explained by the QTL in percentage; PVE (A), proportion of the phenotypic variance explained by additive effect at the current scanning position; PVE (AbyE), eProportion of the phenotypic variance explained by the QTL due to Genotype \times Year interaction; Add, main additive effect; AbyE_01, AbyE_02, and AbyE_03, estimated additive effect of QTL at the current scanning position for each environment (Delhi, Dharwad, and Jodhpur); LeftCI and Right CI, confidence interval calculated by one-LOD drop from the estimated QTL position.

TABLE 9 | List of pearl millet genotypes used to validate the flanked markers of QTLs for grain iron and zinc content.

S. No	Genotype	Fe (mg/kg)	Zn (mg/kg)
1	PPMI 708	114.57	73.68
2	PPMI 1102	111.64	69.77
3	PPMFeZMP 199	110.15	65.95
4	PPMI 1278	101.02	43.50
5	841B	95.86	49.79
6	PPMI 1232	92.44	69.33
7	PPMI 1107	87.31	58.36
8	PPMI 1225	85.03	49.9
9	PPMI 1089	82.66	45.64
10	PPMWGI 146	80.35	32.89
11	PPMI 1231	79.92	35.08
12	PPMI 1011	40.64	40.92
13	PPMI 759	39.98	34.41
14	J2467	39.38	34.93
15	PPMI 1087	39.22	40.26
16	PPMFeZMP 47	38.7	40.46
17	PPMFeZMP 34	36.31	45.10
18	J2405	34.71	37.48
19	H77/833-2	34.03	38.2
20	5540B	28.96	26.62
21	ICMB 98222	98	51.43
22	ICMB 93222	82	61.5
23	ICMB 07999	37	45.12
24	ICMB 92111	40	30

statistically significant divergent phenotypes. The BLUP means of RILs were significant for Fe and Zn content. A very wide range of variation was detected among the population for grain Fe and Zn content (Table 3).

Wider range of variation existed in the present mapping population for grain yield per plant (18.89–36.70 g) and thousand seed weight (6.84–11.79 g). Wide range of variation

was earlier reported by many workers in pearl millet for yield and micronutrient traits (Govindaraj et al., 2011; Sankar et al., 2013; Vinodhana et al., 2013; Bashir et al., 2014; Sabiel et al., 2014; Anuradha et al., 2017). Greater amount of variation in the present population show that there is a lot of scope for selection of better lines to increase the grain yield in pearl millet.

The operational heritability (H^2) estimates were very high and the genotypic components of variance (σ^2g) for both traits were significant. Similar study was also done by Kumar et al. (2016, 2018). The trait heritability information is helpful to choose the selection procedure to be followed to improve the trait. Broad-sense heritability was high enough to allow effective QTL mapping, with moderately high to high proportions of genetic variance for the traits investigated. Higher estimates of heritability with genetic advance as a percent of mean were observed, suggesting the presence of additive gene action efficacy of selection for both traits. The traits which expressed high heritability and low genetic advance showed non additive gene action, hence heterosis breeding would be recommended for these traits. Sumathi et al. (2010) and Govindaraj et al. (2011) were also found similar result.

Frequency Distributions

The frequency distributions of the overall means for grain iron and zinc content in the mapping population are represented by histogram (Figure 1). Likewise, previously Kumar et al. (2016, 2018) were also concluded the normal frequency distribution for both traits in pearl millet. The normal Frequency distribution for both traits was also observed in other crop like wheat (Crespo-Herrera et al., 2017). Both traits in the RIL population have phenotypic normal distributions and transgressive segregations, indicating polygenic inheritance. The presence of transgressive segregation also suggests genetic recombination (Falconer, 1996), which means that both favorable and unfavorable alleles for the traits are dispersed between the parents.

TABLE 10 | Identification of putative candidate gene within QTL involved in enhancing Fe/Zn content.

S. No	QTL	Gene	Swiss sequence ID	Organism	Identities	Expect score	Positive
1	QFe 2.1, QZn 2.1	Ferritin-1, chloroplastic	K3Y9N0	<i>Setaria italica</i>	220/249 (88%)	3e-153	223/249 (89%)
2	QFe 3.1, QZn 3.1	Aluminum-activated malate transporter 5	K3Z099	<i>Setaria italica</i>	439/471 (93%)	0.0	447/471 (94%)
3	QFe 5.1	Potassium transporter 25	K3YQ36	<i>Setaria italica</i>	613/793 (77%)	0.0	630/793 (79%)
4		Probable magnesium transporter NIPA4	K3Z7C3	<i>Setaria italica</i>	76/77 (99%)	8e-40	77/77 (100%)
5		Metal transporter Nramp3 isoform X2	A0A368R0C3	<i>Setaria italica</i>	352/366 (96%)	0.0	353/366 (96%)
6		Potassium transporter 3	A0A1D6N226	<i>Zea mays</i>	188/345 (54%)	9e-120	234/345 (67%)
7		Probable cadmium/zinc-transporting ATPase HMA1, chloroplastic isoform X2	XP_012701051.1	<i>Setaria italica</i>	606/722 (84%)	0.0	626/722 (86%)
8	QFe 7.1	Cadmium/zinc-transporting ATPase HMA2	K3XUY1	<i>Setaria italica</i>	1,020/1,104 (92%)	0.0	1,042/1,104 (94%)
9		Potassium transporter 3	A0A1D6N226	<i>Zea mays</i>	298/520 (57%)	0.0	351/520 (67%)
10		Potassium transporter 3	A0A1D6N226	<i>Zea mays</i>		0.0	466/537 (86%)

Correlation Analysis

The understanding of the association between nutritional and agronomic traits will support the breeders to select suitable selection/breeding program for genetic improvement of associated traits such as mineral content. Several crops have been studied to determine the correlation between grain Fe and Zn contents. Iron and zinc didn't record any significant association with grain yield and thousand seed weight albeit they in turn are correlated. In many previous studies there was negative association of Fe and Zn with yield (Kanatti et al., 2014), which hampers the effective selection of all these traits. This study is in accordance with Anuradha et al. (2017), where they didn't notice any associations. Hence, there is a possibility for selection of high Fe and Zn content along with high yield. In current study, positive correlation between iron and zinc was observed as also reported in many previous studies (Rai et al., 2012; Govindaraj et al., 2013; Kanatti et al., 2014; Kumar et al., 2016, 2018; Anuradha et al., 2017, 2018; Velu et al., 2017). This could be due to common molecular mechanisms that regulate mineral uptake and metabolism in grains, or to common transporters controlling mineral movement within plants (Vreugdenhil et al., 2004; Ghandilyan et al., 2006; Magallanes-López et al., 2017). The strong association between these minerals in populations may be due to the co-segregation of genes for these traits. The direction and intensity of the association suggest that co-transferring superior alleles regulating these traits into the genetic backgrounds of elite lines could provide good opportunities for simultaneous genetic improvement of both micronutrients (Allouis et al., 2001). Genotypic correlations were found to be higher for all traits, implying that there is less interaction between the genetic makeup of the traits and environment.

Linkage Map

Linkage mapping is essential for QTL mapping and marker-assisted breeding programs. The most preferred mapping population is RIL, but the accuracy of the mapping resolution is dependent on population size. The linkage maps have been predominantly constructed using F_2 populations and up to certain extent RILs in pearl millet (Ferreira et al., 2006). Only few previous studies were done on linkage map and QTLs in pearl millet like Supriya et al. (2011) had constructed the linkage map based on 321 loci (258 DArTs and 63 SSRs) for 140 RILs. Similarly, Kumar et al. (2016, 2018) had also used DArTs and SSRs for construction of linkage map and used a smaller number of SSRs for genotyping of RILs and map development. In our study, we found 151 polymorphic SSRs out of 366 and used for linkage map development for 210 RILs. The mapping population size and number of SSRs was high compared to the Supriya et al. (2011) and Kumar et al. (2016, 2018). The present study recorded lesser polymorphism for SSRs compared to that of Senthilvel et al. (2008) who reported 74%. It may be because the parents used here for the development of RILs are closely related than that of Senthilvel et al. (2008). The distribution of SSRs across the linkage group was uniform in the current study. In terms of SSR marker positioning on the genetic map, all SSRs mapped on the similar linkage group (LG) as previously reported pearl millet maps by

various researchers (Yadav et al., 2004; Senthilvel et al., 2008; Supriya et al., 2011; Kumar et al., 2016, 2018). One hundred fifty-one polymorphic SSR markers were assigned to seven previously established pearl millet linkage groups (Figure 2). The present map had a larger length than the one reported by Yadav et al. (2004); Supriya et al. (2011) and Kumar et al. (2016, 2018).

Recombinant inbred lines usually exhibit segregation distortion, because, many recessive lethal genes become homozygous during the RIL development process (Kumar et al., 2016). Segregation distortion was found in 31.57% of the loci on the current linkage map. Whereas Kumar et al. (2018) found 60% segregation distortion of DArTs and SSRs. Pollen abortion is most common in pearl millet than abnormalities in female gametes, leading to a comparatively greater loss of male alleles and the subsequent skewness toward the female parent (Kumar et al., 2018). Distortion from expected Mendelian segregation has been observed previously in maize (Wendel et al., 1987; Lu et al., 2002), barley (Graner et al., 1991; Devaux et al., 1995), rice (Causse et al., 1994; Xu et al., 1997), wheat (Blanco et al., 2004; Quarrie et al., 2005), and pearl millet (Supriya et al., 2011). The protogynous nature of pearl millet also contributes to segregation distortion (Liu et al., 1994). According to Cloutier et al. (1997) and Livingstone et al. (1999), residual heterozygosity and inbreeding depression during inbred development may also contribute to segregation distortion. The residual heterozygosity in some RIL may be advantageous at the same time, because deleterious genetic combinations in the form of reduced fitness or lethality can be avoided. The segregation of nearly all loci of LG 3 was distorted in comparison to earlier reports, which may explain part of the increase in LG 3 map length.

The presence of large gaps between the centromere and telomere is a characteristic feature of pearl millet linkage maps. Senthilvel et al. (2008) found a large gap in LG 4. The previously built framework map for the cross ICMB 841-P3 \times 863B-P2 had big gaps in LG 2 and LG 7 (Senthilvel et al., 2008; Yadav et al., 2011). This new population also revealed a large gap (>25 cM) in LG 1, LG 2, LG 4, and LG 6, possibly due to extreme recombination localization at the ends of LG. The interval distance between the markers in LG 4 was high suggesting additional polymorphic markers are needed to fill-up the gap. According to many researchers like Devos et al. (2000); Varshney et al. (2006), and Senthilvel et al. (2008), large gaps in the distal regions reflect regions of high recombination, rather than a lack of markers in these regions. However, these linkage groups are still incomplete, and need the addition of new markers that are located on the distal regions of the linkage groups. The number of large gaps has tried to decrease in the present study, although there is still the possibility to map more markers to fill these gaps.

Quantitative Trait Loci for Grain Iron (Fe) and Zinc (Zn) Content and Environment Interaction

The use of QTL analysis to map populations is useful not only for identifying genomic regions associated with traits of interest, but also for using the associated marker information in breeding programs to integrate particular loci in elite germplasm. The

application of QTL by environment interaction with composite interval mapping was used in the present study (Li et al., 2015), it was possible to locate several genomic regions associated with GFe and GZn during the three consecutive years over three diverse locations. Our findings support other authors' findings that GFe and GZn are quantitative traits (Kumar et al., 2016; Anuradha et al., 2017). These previous QTL studies have also mapped QTL or genomic region for GFe and GZn in different linkage group of pearl millet, including LG 3 and LG 5, with PVE 13.6 and 10.0%, respectively, for iron content (Kumar et al., 2016). On the other hand, Anuradha et al. (2017), identified different markers on different linkage groups for both high grain iron and zinc content, namely PSMP 2261 (LG 5), IPES0096 (LG 7), IPES 0180 (LG3), and Xsinramp 6 (unmapped). Kumar et al. (2016) reported the co-localization of high grain iron and zinc content alleles/ QTLs in pearl millet. It may happen because starting from uptake to final deposition into grain of iron and zinc may share some common pathways (Grotz and Guerinot, 2006). In present study we found four repetitive QTL on LG 2, LG 3, LG 5, and LG 7 with PVE ranging 16.0–18.6, 14.83–16.26, 16.88–19.63, and 14.59–16.61% in E1, E2, and E3, respectively, and the largest PVE (19.63%) was displayed by *QGF_e.E1.2014–2016_Q3* in LG 5. Kumar et al. (2016) has mapped one SSR marker (IPES142) on LG 3 and Anuradha et al. (2017) has mapped IPES180 on LG3. In the present study one QTL was mapped between IPES142–IPES180 on LG 3 which is in accordance with that of Kumar et al. (2016) and Anuradha et al. (2017). Further, Anuradha et al. (2017) mapped a marker ICMP 3092 on LG 7 for Fe. Similarly, we also identified one repetitive QTL between IPES0206–IPES0015 on LG 7 and interestingly the nearest marker to this QTL is ICMP 3092 as per Rajaram et al. (2013). However, the genomic regions of this study and other studies are not comparable since Kumar et al. (2016) used DArT markers while this study used SSR markers. Similarly, the mapping method used by Anuradha et al. (2017) was association mapping while it is QTL in the present study.

The genotype by environment interaction, particularly the cross-over type interactions, has significant implications for crop performance and breeding. The ideal case for pearl millet biofortification is to obtain stable pearl millet genotypes that perform well without cross-over interaction when tested in different environments or years in a particular geographical area. The analysis of the phenotypic data in our study revealed the presence of a significant $G \times Y$ interaction. QTL analysis revealed that most of the LOD scores for the additive average effect were higher than the LOD score for the interaction (Table 7), showing that QTL with higher LOD (Add) are more stable than those with higher LOD ($G \times Y$) (Li et al., 2015). The additive effect was detected for all the QTLs for grain Fe and Zn, on which the alleles from donor parent had positive additive effect. Gaoh et al. (2020) found the additive \times additive gene effects had the most important effects for grain Fe content, while additive \times dominance gene effects were significant for grain Zn content in pearl millet. Similarly, Garcia-Oliveira et al. (2018) and Kumar et al. (2018) reported that epistasis plays a role in regulating grain iron and zinc content in pearl millet and cereals, respectively. Lu et al. (2008) reported 28 genome-wide additive \times additive interactions

for mineral elements in rice. Anuradha et al. (2012) also reported epistatic interactions between loci on chromosomes 1 and 5 for grain Zn concentration. Jin et al. (2013) reported the additive effect for Zn and partial dominant for Fe in maize.

VALIDATION OF IDENTIFIED QTLs IN A DIVERSE SET OF GERMPLASM

Validation of marker trait associations can reveal their worthiness for further utilization in breeding. Hence, DNA from eleven high, nine low grain iron and four check (two for high and two for low) expressing genotypes were amplified with ten SSR markers which were consistently associated with both the traits. Flanked SSR markers of identified QTLs were compared with the reference Tift₂₃D₂B₁ genome.

In silico Identification of Candidate Gene

The availability of the pearl millet genome sequence has allowed researchers to search for candidate genes involved in Fe and Zn accumulation in grains. The presence of candidate genes was investigated in genomic regions encompassing all QTLs contributing to Fe and Zn in pearl millet. QTLs identified for Fe and Zn content using composite interval mapping were analyzed *In-silico* for the presence of putative candidate genes. In the study, we identified ten metal related genes within the obtained QTLs.

Foods with higher iron content can be produced using modern genetic and molecular technologies. To increase the amount of iron and zinc in wheat, ferritin is over-expressed (Liu et al., 2016). Other mineral concentrations were also affected by ferritin expression, as reported by several studies (Drakakaki et al., 2005; Qu et al., 2005). The promoter selected to control ferritin expression is critical to iron accumulation in specific tissues. The Fe and Zn concentrations in transgenic indica rice grain improved by endosperm-specific glutelin gene promoter of soybean ferritin (Vasconcelos et al., 2003). According to Liu et al. (2016), over-expression of sickle alfalfa ferritin, which is controlled by the seed-storage protein glutelin GluB-1 gene promoter, increases grain Fe and Zn concentrations but also affects mineral homeostasis in transgenic wheat grain. Wheat TaALMT1 (ALMT, for Al-activated Malate Transporter) encoding a malate transporter involved in Al tolerance (Raman et al., 2005). Over expression of TaALMT1 in wheat, barley, and tobacco-cell suspension increases the efflux of Al-activated malate and increases the tolerance to Al stress (Pereira et al., 2010). Root exudation of organic acid such as malate/citrate may play an important role in providing iron to plants (Brown, 1983). Citrate efflux is crucial for iron translocation, and this is mediated by the efflux transporter FRD3 in Arabidopsis (Green and Rogers, 2004) and its orthologue FRDL1 in rice (Yokosho et al., 2016). Barak and Chen (1984) reported a positive effect of K on iron absorption which was associated with acidification of the rhizosphere. Sakaguchi et al. (1999), observed secretion of mugineic acid family phytosiderophores (MAs) from the use of potassium gradient, and content of potassium in barley roots increased with iron-deficiency. Grotz et al. (1998) identified the zinc transporter genes in *Arabidopsis thaliana*. Yang et al. (2020)

studied on activity of zinc transporter (OsZIP9) mediates zinc uptake in rice. Plants acquire Mg from the environment and distribute in the ionic form via Mg^{2+} -permeable transporters. In *Arabidopsis*, MGT6, (plasma membrane-localized magnesium transporter) mediates Mg^{2+} uptake under Mg-limited conditions (Yan et al., 2018). The Magnesium/proton exchanger protein is a vacuolar exchanger of protons with cytosolic Mg^{2+} and Zn^{2+} in *Arabidopsis*. Hence, it may be possible that Mg and Zn share a common pathway of transport in plants.

Ferritin Gene (Ppmsb001–Ppmsb002)

Plant ferritins are important iron storage proteins that serve as both an acceptor and a donor of iron in metabolic processes (Briat et al., 2010). It shares structurally and functionally similarities with animal ferritins. Plant ferritin is generally observed in cells which are inactive in photosynthesis such as seed, root, root nodules, and others. Plant ferritins are nuclear encoded and located primarily in plastids (Zielińska-Dawidziak, 2015). This protein also controls the cellular concentration of transition metals, metal ions apart from iron in the mineral core (including cadmium, beryllium, zinc, aluminum, and lead) (Kumar and Prasad, 1999). Many gene families like FER, ZIP, NRAMP, YSL, and NAS, were involved for mineral translocation from soil to the grain filling site. These gene families have been identified in plants such as *Arabidopsis thaliana* (Weber et al., 2004), *Setaria italica* (Alagarasan et al., 2017), and *Oryza sativa* (Neeraja et al., 2018). In the current report, the ferritin gene was reported by *in silico* method, which is similar to the Ferritin-like (*PglFER1*) gene family previously reported by Mahendrakar et al. (2020). With reference to previous studies it is clear that ferritin gene regulates the iron metabolism in plants.

Al^{3+} Transporter in Plants (IPES0142–IPES0180)

Aluminum is a potentially phytotoxic metal and Al tolerant plants excrete organic acids such as malate, citrate, and oxalate from root tips, depending on the plant (Ryan et al., 2011). The organic acid exudation is vital for plants tolerating metal and nutritional stress at the root-soil interface. Aluminum interferes the uptake or transport of nutrients such as Ca, B, Fe, Mn, P, Mg, and Cu or K (Keltjens and Tan, 1993; Keltjens, 1995; Lukaszewski and Blevins, 1996; Ślaski et al., 1996). Root elongation and nutrient (B, Fe, Mg, Ca, and P) absorption may also be inhibited by Al or low pH soils. Root exudation of organic acid such as malate/citrate may play a key role in supplying Fe to plants (Brown, 1983). Hence, with the consideration of previous studies, we assumed that the aluminum activated malate transporter gene is responsible to cease the aluminum absorption by plant root and also induce the iron absorption.

K^+ Transporters in Plants (IPES0142–IPES0180, IPES0157–IPES0093, and IPES0206–IPES0015)

Oertli and Opoku (1974), as well as Barak and Chen (1984), identified a positive effect of K on Fe nutrition, which was linked to rhizosphere acidification (excessive K^+ uptake and consequent release of H^+ ions by roots to maintain a cation/anion balance). The activity of H^+ ATPase pumps located at the plasma membrane is responsible for rhizosphere acidification, which is

essential for nutrient acquisition. The liberation of proton is in the favor of uptake of Fe and other macro and micronutrients, especially under the deficiency conditions (Houmani et al., 2015). Hughes et al. (1992) found a significant effect of K^+ on proton efflux and ferric reduction mechanisms (Strategy I of iron uptaking by roots). They also identified that K^+ play specific physiological functions in the biosynthetic pathway of mugineic acid production and in the transport of Fe^{3+} -phytosiderophore complex (Strategy II). So, we can assume that K^+ transporter also regulate the iron absorption in plants under iron stress condition.

Zn^{2+} Transporters in Plants (XIPES0157–XIPES0093)

In present study, we have notified one zinc transporter gene of ZIP superfamily that was earlier reported in pearl millet by Mahendrakar et al. (2020). In higher plants, members of the ZIP family are involved in metal uptake, transport, and accumulation of iron and zinc in plant cells. The first identified zinc transporter genes were ZIP1, ZIP2, and ZIP3 in *Arabidopsis thaliana* (Grotz et al., 1998). Several ZIPs, including OsIRT1, OsIRT2, OsZIP1, OsZIP3, OsZIP4, OsZIP5, OsZIP7, and OsZIP8, have been identified as being responsible for Zn uptake from soil, translocation within the root and from the root to the shoot, and storage in rice seeds (Yang et al., 2009; Lee et al., 2010).

Mg^{2+} Transporters in Plants (IPES0157–IPES0093)

Mg^{2+} transporters are used to maintain Mg in plant organs. Among the proteins potentially involved in Mg^{2+} transport, Magnesium transporter (MGT) family Mg^{2+} transporters plant was the most well-investigated protein (Kobayashi and Tanoi, 2015). Meanwhile, the other transporters of Mg^{2+} are also possible such as Non-selective cation channels (NSCCs) (Guo et al., 2003). According to Demidchik and Maathuis (2007), Voltage-independent NSCC (VI-NSCC) reported the uptake of several cations including Mg^{2+} , Ca^{2+} , and Zn^{2+} at the resting membrane potentials. The *Arabidopsis* Magnesium/proton exchanger (MHX) protein is a vacuolar exchanger of protons with cytosolic Mg^{2+} and Zn^{2+} . Hence, it may be possible that Mg and Zn share a common pathway of transport in plants.

We have identified one of the gene in ferritin gene family that was already reported in pearl millet via *insilico* studies and validated by qRT-PCR gene expression (Mahendrakar et al., 2020). The transporter genes like Al, Mg, and K transporter are reporting for the first time in pearl millet. Whereas these transporter gene family were found previously in many different crops such as Al transporters gene in wheat (Li T. et al., 2017) and Rice (Sonoda et al., 2003); Mg transporter gene in Maize (Li et al., 2016, 2018) and *Arabidopsis* (Li H. et al., 2017); K transporter gene in Maize (Zhang et al., 2012) and Barley (Santa-Maria et al., 1997); and Zn transporter gene in *Arabidopsis* (Grotz et al., 1998), rice (Ishimaru et al., 2005; Lee et al., 2010; Neeraja et al., 2018) and Maize (Li et al., 2013).

CONCLUSION

Enhancing the level of grain Fe and Zn content among elite cultivars of pearl millet become the priority area in answering the

nutritional imbalance among poor inhabitation of the semi-arid tropics, especially women and children. Unravelling the genetic principles involved in inheritance of complex traits like grain micronutrient content in pearl millet via, trait-specific mapping enables rapid genetic gains. Huge genetic variation among RIL population observed under the present study can be further used to develop high yielding micronutrient rich cultivar through heterosis breeding. The two co-localized QTLs on LG 2 and LG 3, for both the minerals found to promising as it could simultaneously transferred to the parental lines of elite pearl millet hybrids through marker-assisted breeding programs.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

CS, SPS, and AK designed and supervised the overall research and contributed in the preparation of the manuscript. CS and SPS provided technical guidance for creating liaison among multi-environment locations, and edited the manuscript for final submission. TS, SMS, JB, and MM executed the field experiments. TS and SMS performed the phenotyping, carried out statistical analysis, and prepared the manuscript draft. CB and NA reviewed the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

- Aboul-Maaty, N. A. F., and Oraby, H. A. S. (2019). Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bull. Natl. Res. Cent.* 43:25. doi: 10.1186/s42269-019-0066-1
- Alagarasan, G., Dubey, M., Aswathy, K. S., and Chandel, G. (2017). Genome wide identification of orthologous ZIP genes associated with zinc and Iron translocation in *Setaria italica*. *Front. Plant Sci.* 8:775. doi: 10.3389/fpls.2017.00775
- Allouis, S., Qi, X., Lindup, S., Gale, M. D., and Devos, K. M. (2001). Construction of a BAC library of pearl millet, *Pennisetum glaucum*. *Theor. Appl. Genet.* 102, 1200–1205. doi: 10.1007/s001220100559
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Anuradha, K., Agarwal, S., Rao, Y. V., Rao, K. V., Viraktamath, B. C., and Sarla, N. (2012). Mapping QTLs and candidate genes for iron and zinc concentrations in unpolished rice of Madhukar × Swarna RILs. *Gene* 508, 233–240. doi: 10.1016/j.gene.2012.07.054
- Anuradha, N., Satyavathi, C. T., Bharadwaj, C., Nepolean, T., Sankar, S. M., Singh, S. P., et al. (2017). Deciphering genomic regions for high grain iron and zinc content using association mapping in pearl millet. *Front. Plant Sci.* 8:412. doi: 10.3389/fpls.2017.00412
- Anuradha, N., Satyavathi, C. T., Bharadwaj, C., Sankar, M., and Pathy, L. (2018). Association of agronomic traits and micronutrients in pearl millet. *Int. J. Chem. Stud.* 6, 181–184.

FUNDING

Financial support of ICAR through the project entitled “CRP on biofortification pearl millet [*Pennisetum glaucum* (L.) Br.]” (ICAR-Sanction number: CS.11/6/2014-IA-IV).

ACKNOWLEDGMENTS

The authors are thankful to A. K. Singh, Director, IARI, New Delhi, Division of Genetics, IARI, New Delhi, ICAR-IARI for providing the necessary facilities for smooth conductance of research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.659789/full#supplementary-material>

Supplementary Figure 1 | A to D shown the QTLs for grain Fe content on LG 2,3,5, and 7 at Delhi location.

Supplementary Figure 2 | A and B shown the QTLs for grain Zn content at LG 2 and 3 at Delhi location.

Supplementary Figure 3 | A to D shown the QTLs for grain Fe content on LG 2,3,5, and 7 at Dharwad location.

Supplementary Figure 4 | A and B shown the QTLs for grain Zn content at LG 2 and 3 at Dharwad location.

Supplementary Figure 5 | A to D shown the QTLs for grain Fe content on LG 2,3,5, and 7 at Jodhpur location.

Supplementary Figure 6 | A and B shown the QTLs for grain Zn content at LG 2 and 3 at Jodhpur location.

- Athar, T., Khan, M. K., Pandey, A., Yilmaz, F. G., Hamurcu, M., Hakki, E. E., et al. (2020). Biofortification and the involved modern approaches. *J. Elem.* 25, 717–731.
- Barak, P., and Chen, Y. (1984). The effect of potassium on iron chlorosis in calcareous soils. *J. Plant Nutr.* 7, 125–133. doi: 10.1080/01904168409363180
- Bashir, E. M., Ali, A. M., Ali, A. M., Melchinger, A. E., Parzies, H. K., and Haussmann, B. I. (2014). Characterization of Sudanese pearl millet germplasm for agro-morphological traits and grain nutritional values. *Plant Genet. Resour.* 12:35. doi: 10.1017/s1479262113000233
- Blanco, A., Simeone, R., Cenci, A., Gadaleta, A., Tanzarella, O. A., Porceddu, E., et al. (2004). Extension of the Messapia × dicoccoides linkage map of *Triticum turgidum* (L.) Thell. *Cell Mol. Biol. Lett.* 9, 529–542.
- Briat, J. F., Duc, C., Ravet, K., and Gaymard, F. (2010). Ferritins and iron storage in plants. *Biochim. Biophys. Acta* 1800, 806–814. doi: 10.1016/j.bbagen.2009.12.003
- Brown, D. J. A. (1983). Effect of calcium and aluminum concentrations on the survival of brown trout (*Salmo trutta*) at low pH. *Bull. Environ. Contam. Toxicol.* 30, 582–587. doi: 10.1007/bf01610178
- Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K., et al. (1994). Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics* 138, 1251–1274. doi: 10.1093/genetics/138.4.1251
- Cloutier, S., Cappadocia, M., and Landry, B. S. (1997). Analysis of RFLP mapping inaccuracy in *Brassica napus* L. *Theor. Appl. Genet.* 95, 83–91. doi: 10.1007/s001220050535

- Colangelo, E. P., and Guerinot, M. L. (2006). Put the metal to the petal: metal uptake and transport throughout plants. *Curr. Opin. Plant Biol.* 9, 322–330. doi: 10.1016/j.pbi.2006.03.015
- Connorton, J. M., Balk, J., and Rodríguez-Celma, J. (2017). Iron homeostasis in plants—a brief overview. *Metallomics* 9, 813–823. doi: 10.1039/c7mt00136c
- Crespo-Herrera, L. A., Govindan, V., Stangoulis, J., Hao, Y., and Singh, R. P. (2017). QTL mapping of grain Zn and Fe concentrations in two hexaploid wheat RIL populations with ample transgressive segregation. *Front. Plant Sci.* 8:1800. doi: 10.3389/fpls.2017.01800
- Demidchik, V., and Maathuis, F. J. (2007). Physiological roles of nonselective cation channels in plants: from salt stress to signalling and development. *New Phytol.* 175, 387–404. doi: 10.1111/j.1469-8137.2007.02128.x
- Devaux, P., Kilian, A., and Kleinhofs, A. (1995). Comparative mapping of the barley genome with male and female recombination-derived, doubled haploid populations. *Mol. Gen. Genet.* 249, 600–608. doi: 10.1007/bf00418029
- Devos, K. M., Pittaway, T. S., Reynolds, A., and Gale, M. D. (2000). Comparative mapping reveals a complex relationship between the pearl millet genome and those of foxtail millet and rice. *Theor. Appl. Genet.* 100, 190–198. doi: 10.1007/s001220050026
- Dhyan, S., Chhonkar, P. K., and Dwivedi, B. S. (2005). *Manual on Soil, Plant and Water Analysis*. New Delhi: Westville Publishing House.
- Doerge, R. W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.* 3, 43–52. doi: 10.1038/nrg703
- Drakakaki, G., Marcel, S., Glahn, R. P., Lund, E. K., Pariagh, S., Fischer, R., et al. (2005). Endosperm-specific co-expression of recombinant soybean ferritin and Aspergillus phytase in maize results in significant increases in the levels of bioavailable iron. *Plant Mol. Biol.* 59, 869–880. doi: 10.1007/s11103-005-1537-3
- Falconer, D. S. (1996). *Introduction to Quantitative Genetics*. Noida: Pearson Education India.
- Ferreira, A., Silva, M. F. D., and Cruz, C. D. (2006). Estimating the effects of population size and type on the accuracy of genetic maps. *Genet. Mol. Biol.* 29, 187–192. doi: 10.1590/s1415-47572006000100033
- Gangashetty, P. I., Motagi, B. N., Pavan, R., and Roodagi, M. B. (2016). “Breeding crop plants for improved human nutrition through biofortification: progress and prospects,” in *Advances in Plant Breeding Strategies: Agronomic, Abiotic and Biotic Stress Traits*, eds J. Al-Khayri, S. Jain, and D. Johnson (Cham: Springer), 35–76. doi: 10.1007/978-3-319-22518-0_2
- Gaoh, B. S. B., Gangashetty, P. I., Mohammed, R., Dzidzienyo, D. K., and Tongoono, P. (2020). Generation mean analysis of pearl millet [*Pennisetum glaucum* (L.) R. Br.] grain iron and zinc contents and agronomic traits in West Africa. *J. Cereal Sci.* 96:103066. doi: 10.1016/j.jcs.2020.103066
- Garcia-Oliveira, A. L., Chander, S., Ortiz, R., Menkir, A., and Gedil, M. (2018). Genetic basis and breeding perspectives of grain iron and zinc enrichment in cereals. *Front. Plant Sci.* 9:937. doi: 10.3389/fpls.2018.00937
- Ghandilyan, A., Vreugdenhil, D., and Aarts, M. G. (2006). Progress in the genetic understanding of plant iron and zinc nutrition. *Physiol. Plant.* 126, 407–417. doi: 10.1111/j.1399-3054.2006.00646.x
- Govindaraj, M., Rai, K. N., and Shanmugasundaram, P. (2016). Intra-population genetic variance for grain iron and zinc contents and agronomic traits in pearl millet. *Crop J.* 4, 48–54. doi: 10.1016/j.cj.2015.11.002
- Govindaraj, M., Rai, K. N., Shanmugasundaram, P., Dwivedi, S. L., Sahrawat, K. L., Muthaiah, A. R., et al. (2013). Combining ability and heterosis for grain iron and zinc densities in pearl millet. *Crop Sci.* 53, 507–517. doi: 10.2135/cropsci2012.08.0477
- Govindaraj, M., Selvi, B., Rajarathinam, S., and Sumathi, P. (2011). Genetic variability and heritability of grain yield components and grain mineral concentration in India's pearl millet (*Pennisetum glaucum* (L.) R. Br.) accessions. *Afr. J. Food Agric. Nutr. Dev.* 11:4758.
- Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., et al. (1991). Construction of an RFLP map of barley. *Theor. Appl. Genet.* 83, 250–256. doi: 10.1007/bf00226259
- Green, L. S., and Rogers, E. E. (2004). FRD3 controls iron localization in *Arabidopsis*. *Plant Physiol.* 136, 2523–2531. doi: 10.1104/pp.104.045633
- Grotz, N., Fox, T., Connolly, E., Park, W., Guerinot, M. L., and Eide, D. (1998). Identification of a family of zinc transporter genes from *Arabidopsis* that respond to zinc deficiency. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7220–7224. doi: 10.1073/pnas.95.12.7220
- Grotz, N., and Guerinot, M. L. (2006). Molecular aspects of Cu, Fe and Zn homeostasis in plants. *Biochim. Biophys. Acta Mol. Cell Res.* 1763, 595–608. doi: 10.1016/j.bbamcr.2006.05.014
- Guo, T. R., Zhang, G. P., Lu, W. Y., Wu, H. P., Chen, J. X., Wu, F. B., et al. (2003). Effect of Al on dry matter accumulation and Al and nutrition contents of barleys differing in Al tolerance. *Plant Nutr. Fertil. Sci.* 9, 324–330.
- Gupta, H. S., Agrawal, P. K., Mahajan, V., Bisht, G. S., Kumar, A., Verma, P., et al. (2009). Quality protein maize for nutritional security: rapid development of short duration hybrids through molecular marker assisted breeding. *Curr. Sci.* 96, 230–237.
- HarvestPlus (2014). *Biofortification Progress Briefs*. Available online at: www.HarvestPlus.org (accessed August, 2014)
- Houmani, H., Rabhi, M., Abdelly, C., and Debez, A. (2015). “Implication of rhizosphere acidification in nutrient uptake by plants: cases of potassium (K), phosphorus (P), and iron (Fe),” in *Crop Production and Global Environmental Issues*, ed. K. Hakeem (Cham: Springer), 103–122. doi: 10.1007/978-3-319-23162-4_4
- Hughes, D. F., Jolley, V. D., and Brown, J. C. (1992). Roles for potassium in the iron-stress response mechanisms of strategy I and strategy II plants. *J. Plant Nutr.* 15, 1821–1839. doi: 10.1080/01904169209364442
- Ishimaru, Y., Suzuki, M., Kobayashi, T., Takahashi, M., Nakanishi, H., Mori, S., et al. (2005). OsZIP4, a novel zinc-regulated zinc transporter in rice. *J. Exp. Bot.* 56, 3207–3214. doi: 10.1093/jxb/eri317
- Jin, T., Zhou, J., Chen, J., Zhu, L., Zhao, Y., and Huang, Y. (2013). The genetic architecture of zinc and iron content in maize grains as revealed by QTL mapping and meta-analysis. *Breed. Sci.* 63, 317–324. doi: 10.1270/jsbbs.63.317
- Kanatti, A., Rai, K. N., Radhika, K., Govindaraj, M., Sahrawat, K. L., Srinivasu, K., et al. (2014). Relationship of grain iron and zinc content with grain yield in pearl millet hybrids. *Crop. Improv.* 41, 91–96.
- Keltjens, W. G. (1995). Magnesium uptake by Al-stressed maize plants with special emphasis on cation interactions at root exchange sites. *Plant Soil* 171, 141–146. doi: 10.1007/bf00009577
- Keltjens, W. G., and Tan, K. (1993). “Interactions between aluminium, magnesium and calcium with different monocotyledonous and dicotyledonous plant species,” in *Plant Nutrition—from Genetic Engineering to Field Practice*, ed. N. J. Barrow (Dordrecht: Springer), 719–722. doi: 10.1007/978-94-011-1880-4_158
- Kobayashi, N. I., and Tanoi, K. (2015). Critical issues in the study of magnesium transport systems and magnesium deficiency symptoms in plants. *Int. J. Mol. Sci.* 16, 23076–23093. doi: 10.3390/ijms160923076
- Kotla, A., Phuke, R., Hariprasanna, K., Mehre, S. P., Rathore, A., Gorthy, S., et al. (2019). Identification of QTLs and candidate genes for high grain Fe and Zn concentration in sorghum [*Sorghum bicolor* (L.) Moench]. *J. Cereal Sci.* 90, 102850. doi: 10.1016/j.jcs.2019.102850
- Kumar, S., Hash, C. T., Nepolean, T., Mahendrakar, M. D., Satyavathi, C. T., Singh, G., et al. (2018). Mapping grain iron and zinc content quantitative trait loci in an inbred-derived immortal population of pearl millet. *Genes* 9:248. doi: 10.3390/genes9050248
- Kumar, S., Hash, C. T., Thirunavukkarasu, N., Singh, G., Rajaram, V., Rathore, A., et al. (2016). Mapping quantitative trait loci controlling high iron and zinc content in self and open pollinated grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 7:1636. doi: 10.3389/fpls.2016.01636
- Kumar, T. R., and Prasad, M. N. V. (1999). Ferritin induction by iron mediated oxidative stress and ABA in *Vigna mungo* (L.) Hepper seedlings: role of antioxidants and free radical scavengers. *J. Plant Physiol.* 155, 652–655. doi: 10.1016/s0176-1617(99)80068-6
- Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E., et al. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1, 174–181. doi: 10.1016/0888-7543(87)90010-3
- Lee, S., Jeong, H. J., Kim, S. A., Lee, J., Guerinot, M. L., and An, G. (2010). OsZIP5 is a plasma membrane zinc transporter in rice. *Plant Mol. Biol.* 73, 507–517. doi: 10.1007/s11103-010-9637-0
- Li, H., Du, H., Huang, K., Chen, X., Liu, T., Gao, S., et al. (2016). Identification, and functional and expression analyses of the CorA/MRS2/MGT-type magnesium transporter family in maize. *Plant Cell Physiol.* 57, 1153–1168. doi: 10.1093/pcp/pcw064

- Li, H., Liu, C., Zhou, L., Zhao, Z., Li, Y., Qu, M., et al. (2018). Molecular and functional characterization of the magnesium transporter gene *ZmMGT12* in maize. *Gene* 665, 167–173. doi: 10.1016/j.gene.2018.04.068
- Li, H., Wang, N., Ding, J., Liu, C., Du, H., Huang, K., et al. (2017). The maize CorA/MRS2/MGT-type Mg transporter, *ZmMGT10*, responses to magnesium deficiency and confers low magnesium tolerance in transgenic *Arabidopsis*. *Plant Mol. Biol.* 95, 269–278. doi: 10.1007/s11103-017-0645-1
- Li, S., Wang, J., and Zhang, L. (2015). Inclusive composite interval mapping of QTL by environment interactions in biparental populations. *PLoS One* 10:e0132414. doi: 10.1371/journal.pone.0132414
- Li, S., Zhou, X., Huang, Y., Zhu, L., Zhang, S., Zhao, Y., et al. (2013). Identification and characterization of the zinc-regulated transporters, iron-regulated transporter-like protein (ZIP) gene family in maize. *BMC Plant Biol.* 13:114. doi: 10.1186/1471-2229-13-114
- Li, T., Liao, K., Xu, X., Gao, Y., Wang, Z., Zhu, X., et al. (2017). Wheat ammonium transporter (AMT) gene family: diversity and possible role in host–pathogen interaction with stem rust. *Front. Plant Sci.* 8:1637. doi: 10.3389/fpls.2017.01637
- Liu, C. J., Witcombe, J. R., Pittaway, T. S., Nash, M., Hash, C. T., Busso, C. S., et al. (1994). An RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). *Theor. Appl. Genet.* 89, 481–487. doi: 10.1007/bf00225384
- Liu, D. J., Wang, Y. B., Guo, C. H., Cong, Q., Gong, X., and Zhang, H. J. (2016). Enhanced iron and zinc accumulation in genetically engineered wheat plants using sickle alfalfa (*Medicago falcata* L.) ferritin gene. *Cereal Res. Commun.* 44, 24–34. doi: 10.1556/0806.43.2015.039
- Livingstone, K. D., Lackney, V. K., Blauth, J. R., Van Wijk, R. I. K., and Jahn, M. K. (1999). Genome mapping in Capsicum and the evolution of genome structure in the Solanaceae. *Genetics* 152, 1183–1202.
- Lu, H., Romero-Severson, J., and Bernardo, R. (2002). Chromosomal regions associated with segregation distortion in maize. *Theor. Appl. Genet.* 105, 622–628. doi: 10.1007/s00122-002-0970-9
- Lu, K., Li, L., Zheng, X., Zhang, Z., Mou, T., and Hu, Z. (2008). Quantitative trait loci controlling Cu, Ca, Zn, Mn and Fe content in rice grains. *J. Genet.* 87:305. doi: 10.1007/s12041-008-0049-8
- Luciano Da Costa, E. S., Wang, S., and Zeng, Z. B. (2012). Composite interval mapping and multiple interval mapping: procedures and guidelines for using Windows QTL Cartographer. *Methods Mol. Biol.* 871, 75–119. doi: 10.1007/978-1-61779-785-9_6
- Lukaszewski, K. M., and Blevins, D. G. (1996). Root growth inhibition in boron-deficient or aluminum-stressed squash may be a result of impaired ascorbate metabolism. *Plant Physiol.* 112, 1135–1140. doi: 10.1104/pp.112.3.1135
- Magallanes-López, A. M., Hernandez-Espinosa, N., Velu, G., Posadas-Romano, G., Ordoñez-Villegas, V. M. G., Crossa, J., et al. (2017). Variability in iron, zinc and phytic acid content in a worldwide collection of commercial durum wheat cultivars and the effect of reduced irrigation on these traits. *Food Chem.* 237, 499–505. doi: 10.1016/j.foodchem.2017.05.110
- Mahendrakar, M. D., Parveda, M., Kishor, P. K., and Srivastava, R. K. (2020). Discovery and validation of candidate genes for grain iron and zinc metabolism in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Sci. Rep.* 10:16562.
- Mohan, M., Nair, S., Bhagwat, A., Krishna, T. G., Yano, M., Bhatia, C. R., et al. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol. Breed.* 3, 87–103.
- Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Choudhary, M., Saha, S., Bhat, J. S., et al. (2014). Development of β -carotene rich maize hybrids through marker-assisted introgression of β -carotene hydroxylase allele. *PLoS One* 9:e113583. doi: 10.1371/journal.pone.0113583
- Neeraja, C. N., Kulkarni, K. S., Madhu Babu, P., Sanjeeva Rao, D., Surekha, K., and Ravindra Babu, V. (2018). Transporter genes identified in landraces associated with high zinc in polished rice through panicle transcriptome for biofortification. *PLoS One* 13:e0192362. doi: 10.1371/journal.pone.0192362
- Nozoye, T., Nagasaka, S., Kobayashi, T., Takahashi, M., Sato, Y., Sato, Y., et al. (2011). Phytosiderophore efflux transporters are crucial for iron acquisition in graminaceous plants. *J. Biol. Chem.* 286, 5446–5454. doi: 10.1074/jbc.m110.180026
- Oertli, J. J., and Opoku, A. A. (1974). Interaction of potassium in the availability and uptake of iron from ferric hydroxide. *Soil Sci. Soc. Am. J.* 38, 451–454. doi: 10.2136/sssaj1974.03615995003800030023x
- Patterson, H. D., and Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika* 63, 83–92. doi: 10.1093/biomet/63.1.83
- Pereira, J. F., Zhou, G., Delhaize, E., Richardson, T., Zhou, M., and Ryan, P. R. (2010). Engineering greater aluminium resistance in wheat by over-expressing TaALMT1. *Ann. Bot.* 106, 205–214. doi: 10.1093/aob/mcq058
- Pfeiffer, W. H., and McClafferty, B. (2007). HarvestPlus: breeding crops for better nutrition. *Crop Sci.* 47, S-88–S-105.
- Qu, L., Yoshihara, T., Ooyama, A., Goto, F., and Takaiwa, F. (2005). Iron accumulation does not parallel the high expression level of ferritin in transgenic rice seeds. *Planta* 222, 225–233. doi: 10.1007/s00425-005-1530-8
- Quarrie, S. A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., et al. (2005). A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring \times SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* 110, 865–880. doi: 10.1007/s00122-004-1902-7
- Rai, K. N., Govindaraj, M., and Rao, A. S. (2012). Genetic enhancement of grain iron and zinc content in pearl millet. *Qual. Assur. Saf. Crop. Foods* 4, 119–125. doi: 10.1111/j.1757-837x.2012.00135.x
- Rai, K. N., Velu, G., Govindaraj, M., Upadhyaya, H. D., Rao, A. S., Shivade, H., et al. (2015). Iniadi pearl millet germplasm as a valuable genetic resource for high grain iron and zinc densities. *Plant Genet. Resour.* 13, 1–8.
- Rajaram, V., Nepolean, T., Senthilvel, S., Varshney, R. K., Vadez, V., Srivastava, R. K., et al. (2013). Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. *BMC Genomics* 14:159. doi: 10.1186/1471-2164-14-159
- Raman, H., Zhang, K., Cakir, M., Appels, R., Garvin, D. F., Maron, L. G., et al. (2005). Molecular characterization and mapping of ALMT1, the aluminium-tolerance gene of bread wheat (*Triticum aestivum* L.). *Genome* 48, 781–791. doi: 10.1139/g05-054
- Roberts, L. A., Pierson, A. J., Panaviene, Z., and Walker, E. L. (2004). Yellow stripe1. Expanded roles for the maize iron-phytosiderophore transporter. *Plant Physiol.* 135, 112–120. doi: 10.1104/pp.103.037572
- Ryan, P. R., Tyerman, S. D., Sasaki, T., Furuichi, T., Yamamoto, Y., Zhang, W. H., et al. (2011). The identification of aluminium-resistance genes provides opportunities for enhancing crop production on acid soils. *J. Exp. Bot.* 62, 9–20. doi: 10.1093/jxb/erq272
- Sabiel, S. A., Ismail, M. I., Abdalla, E., Osman, K. A., and Ali, A. M. (2014). Genetic variation among pearl millet genotypes for yield and its components in semi-arid zone Sudan. *IJACS* 7, 822–826.
- Sakaguchi, T., Nishizawa, N. K., Nakanishi, H., Yoshimura, E., and Mori, S. (1999). The role of potassium in the secretion of mugineic acids family phytosiderophores from iron-deficient barley roots. *Plant Soil* 215, 221–227.
- Salamov, A. A., and Solov'yev, V. V. (2000). Ab initio gene finding in *Drosophila* genomic DNA. *Genome Res.* 10, 516–522. doi: 10.1101/gr.10.4.516
- Sankar, S. M., Satyavathi, C. T., Singh, M. P., Bharadwaj, C., Singh, S. P., and Barthakur, S. (2013). Genetic variability and association studies in pearl millet for grain yield and high temperature stress tolerance. *Indian J. Dryland Agric. Res. Dev.* 28, 71–76.
- Santa-Maria, G. E., Rubio, F., Dubcovsky, J., and Rodríguez-Navarro, A. (1997). The HAK1 gene of barley is a member of a large gene family and encodes a high-affinity potassium transporter. *Plant Cell* 9, 2281–2289. doi: 10.2307/3870585
- Satyavathi, C. T., Sankar, S. M., Singh, S. P., Bhowmick, P., Bhat, J., Singh, O., et al. (2015). Stability analysis of grain iron and zinc content in pearl millet (*Pennisetum glaucum* (L.) R. Br.). *J. Trop. Agric.* 33, 1387–1394.
- Senthilvel, S., Jayashree, B., Mahalakshmi, V., Kumar, P. S., Nakka, S., Nepolean, T., et al. (2008). Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. *BMC Plant Biol.* 8:119. doi: 10.1186/1471-2229-8-119
- Shariatipour, N., and Heidari, B. (2020). “Genetic-based biofortification of staple food crops to meet zinc and iron deficiency-related challenges,” in *Plant Micronutrients*, eds T. Aftab and K. R. Hakeem (Cham: Springer), 173–223. doi: 10.1007/978-3-030-49856-6_8
- Shivhare, R., and Lata, C. (2016). Selection of suitable reference genes for assessing gene expression in pearl millet under different abiotic stresses and their combinations. *Sci. Rep.* 6:23036.
- Singh, V. K., Singh, A., Singh, S. P., Ellur, R. K., Choudhary, V., Sarkel, S., et al. (2012). Incorporation of blast resistance into “PRR78”, an elite Basmati rice

- restorer line, through marker assisted backcross breeding. *Field Crops Res.* 128, 8–16. doi: 10.1016/j.fcr.2011.12.003
- Ślaski, J. J., Zhang, G., Basu, U., Stephens, J. L., and Taylor, G. J. (1996). Aluminum resistance in wheat (*Triticum aestivum*) is associated with rapid, Al-induced changes in activities of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in root apices. *Physiol. Plant.* 98, 477–484. doi: 10.1034/j.1399-3054.1996.980307.x
- Sonoda, Y., Ikeda, A., Saiki, S., Yamaya, T., and Yamaguchi, J. (2003). Feedback regulation of the ammonium transporter gene family AMT1 by glutamine in rice. *Plant Cell Physiol* 44, 1396–1402. doi: 10.1093/pcp/pcg169
- Sumathi, P., Madineni, S., and Veerabhadhiran, P. (2010). Genetic variability for different biometrical traits in pearl millet genotypes (*Pennisetum glaucum* LR BR.). *Electron. J. Plant Breed.* 1, 437–440.
- Supriya, A., Senthilvel, S., Nepolean, T., Eshwar, K., Rajaram, V., Shaw, R., et al. (2011). Development of a molecular linkage map of pearl millet integrating DArT and SSR markers. *Theor. Appl. Genet.* 123, 239–250. doi: 10.1007/s00122-011-1580-1
- Varshney, R. K., Hoisington, D. A., and Tyagi, A. K. (2006). Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol.* 24, 490–499. doi: 10.1016/j.tibtech.2006.08.006
- Vasconcelos, M., Datta, K., Oliva, N., Khalekuzzaman, M., Torrizo, L., Krishnan, S., et al. (2003). Enhanced iron and zinc accumulation in transgenic rice with the ferritin gene. *Plant Sci.* 164, 371–378. doi: 10.1016/s0168-9452(02)00421-1
- Velu, G., Tulus, Y., Gomez-Becerra, H. F., Hao, Y., Demir, L., Kara, R., et al. (2017). QTL mapping for grain zinc and iron concentrations and zinc efficiency in a tetraploid and hexaploid wheat mapping populations. *Plant Soil.* 411, 81–99. doi: 10.1007/s11104-016-3025-8
- Vinodhana, N. K., Sumathi, P., and Sathya, M. (2013). Genetic variability and inter-relationship among morpho-economic traits of pearl millet (*Pennisetum glaucum* (L.) r. Br.) and their implications in selection. *IJPAES* 3, 145–149.
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77
- Vreugdenhil, D., Aarts, M. G. M., Koornneef, M., Nelissen, H., and Ernst, W. H. O. (2004). Natural variation and QTL analysis for cationic mineral content in seeds of *Arabidopsis thaliana*. *Plant Cell Environ.* 27, 828–839. doi: 10.1111/j.1365-3040.2004.01189.x
- Wang, X., Yang, S., Chen, Y., Zhang, S., Zhao, Q., Li, M., et al. (2018). Comparative genome-wide characterization leading to simple sequence repeat marker development for Nicotiana. *BMC Genomics* 19:500. doi: 10.1186/s12864-018-4878-4
- Weber, M., Harada, E., Vess, C., Roepenack-Lahaye, E. V., and Clemens, S. (2004). Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *Plant J.* 37, 269–281. doi: 10.1046/j.1365-3113x.2003.01960.x
- Wendel, J. F., Edwards, M. D., and Stuber, C. W. (1987). Evidence for multilocus genetic control of preferential fertilisation in maize. *Heredity* 58, 297–301. doi: 10.1038/hdy.1987.44
- Xu, Y., Zhu, L., Xiao, J., Huang, N., and McCouch, S. R. (1997). Chromosomal regions associated with segregation distortion of molecular markers in F₂, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). *Mol. Gen. Genet.* 253, 535–545. doi: 10.1007/s004380050355
- Yadav, R. S., Hash, C. T., Bidingier, F. R., Devos, K. M., and Howarth, C. J. (2004). Genomic regions associated with grain yield and aspects of post-flowering drought tolerance in pearl millet across stress environments and tester background. *Euphytica* 136, 265–277. doi: 10.1023/b:euph.0000032711.34599.3a
- Yadav, R. S., Sehgal, D., and Vadez, V. (2011). Using genetic mapping and genomics approaches in understanding and improving drought tolerance in pearl millet. *J. Exp. Bot.* 62, 397–408. doi: 10.1093/jxb/erq265
- Yan, Y. W., Mao, D. D., Yang, L., Qi, J. L., Zhang, X. X., Tang, Q. L., et al. (2018). Magnesium transporter MGT6 plays an essential role in maintaining magnesium homeostasis and regulating high magnesium tolerance in *Arabidopsis*. *Front. Plant Sci.* 9:274. doi: 10.3389/fpls.2018.00274
- Yang, M., Li, Y., Liu, Z., Tian, J., Liang, L., Qiu, Y., et al. (2020). A high activity zinc transporter OsZIP9 mediates zinc uptake in rice. *Plant J.* 103, 1695–1709. doi: 10.1111/tpl.14855
- Yang, Z., Wu, Y., Li, Y., Ling, H. Q., and Chu, C. (2009). OsMT1a, a type 1 metallothionein, plays the pivotal role in zinc homeostasis and drought tolerance in rice. *Plant Mol. Biol.* 70, 219–229. doi: 10.1007/s11103-009-9466-1
- Yokosho, K., Yamaji, N., and Ma, J. F. (2016). OsFRDL1 expressed in nodes is required for distribution of iron to grains in rice. *J. Exp. Bot.* 67, 5485–5494. doi: 10.1093/jxb/erw314
- Zhang, Z., Zhang, J., Chen, Y., Li, R., Wang, H., and Wei, J. (2012). Genome-wide analysis and identification of HAK potassium transporter gene family in maize (*Zeamays* L.). *Mol. Biol. Rep.* 39, 8465–8473. doi: 10.1007/s11033-012-1700-2
- Zielińska-Dawidziak, M. (2015). Plant ferritin—a source of iron to prevent its deficiency. *Nutrients* 7, 1184–1201. doi: 10.3390/nu7021184

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Singhal, Satyavathi, Singh, Kumar, Sankar, Bhardwaj, Mallik, Bhat, Anuradha and Singh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Sorghum Grain Mold Disease Complex: Pathogens, Host Responses, and the Bioactive Metabolites at Play

Arlyn Ackerman¹, Anthony Wenndt² and Richard Boyles^{1*}

¹ Cereal Grains Breeding and Genetics, Pee Dee Research and Education Center, Department of Plant & Environmental Sciences, Clemson University, Florence, SC, United States, ² Plant Pathology and Plant-Microbe Biology, The School of Integrated Plant Sciences, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:

Mahalingam Govindaraj,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

Reviewed by:

Yi Ding,
The University of Sydney, Australia
Raju Ghosh,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

*Correspondence:

Richard Boyles
rboyles@clemson.edu

Specialty section:

This article was submitted to
Plant Pathogen Interactions,
a section of the journal
Frontiers in Plant Science

Received: 28 January 2021

Accepted: 21 April 2021

Published: 28 May 2021

Citation:

Ackerman A, Wenndt A and
Boyles R (2021) The Sorghum Grain
Mold Disease Complex: Pathogens,
Host Responses, and the Bioactive
Metabolites at Play.
Front. Plant Sci. 12:660171.
doi: 10.3389/fpls.2021.660171

Grain mold is a major concern in sorghum [*Sorghum bicolor* (L.) Moench] production systems, threatening grain quality, safety, and nutritional value as both human food and livestock feed. The crop's nutritional value, environmental resilience, and economic promise poise sorghum for increased acreage, especially in light of the growing pressures of climate change on global food systems. In order to fully take advantage of this potential, sorghum improvement efforts and production systems must be proactive in managing the sorghum grain mold disease complex, which not only jeopardizes agricultural productivity and profitability, but is also the culprit of harmful mycotoxins that warrant substantial public health concern. The robust scholarly literature from the 1980s to the early 2000s yielded valuable insights and key comprehensive reviews of the grain mold disease complex. Nevertheless, there remains a substantial gap in understanding the complex multi-organismal dynamics that underpin the plant-pathogen interactions involved – a gap that must be filled in order to deliver improved germplasm that is not only capable of withstanding the pressures of climate change, but also wields robust resistance to disease and mycotoxin accumulation. The present review seeks to provide an updated perspective of the sorghum grain mold disease complex, bolstered by recent advances in the understanding of the genetic and the biochemical interactions among the fungal pathogens, their corresponding mycotoxins, and the sorghum host. Critical components of the sorghum grain mold disease complex are summarized in narrative format to consolidate a collection of important concepts: (1) the current state of sorghum grain mold in research and production systems; (2) overview of the individual pathogens that contribute to the grain mold complex; (3) the mycotoxin-producing potential of these pathogens on sorghum and other substrates; and (4) a systems biology approach to the understanding of host responses.

Keywords: sorghum grain mold, mycotoxins, disease resistance, host-pathogen interactions, phenolics, phenylpropanoid pathway, flavonoid pathway

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is grown internationally for its robust ability to withstand harsh climates, underlined by great water use efficiency and a diverse biochemical profile boasting high antioxidant potential. Sorghum exhibits robust environmental resilience and yield stability while being capable of reaching potential yield levels at lower input costs than other major cereals. These desirable traits have the crop posed for increased acreage. However, in many areas of agronomic sorghum growth such as regions in the tropical and subtropical climates, high heat and extended periods of humid conditions (relative humidity > 70%) are common. Regions such as these demonstrate inherently warm air capable of holding increased levels of water vapor, and will become more common and expansive as climate change occurs (Willett et al., 2014). Unfortunately, high heat and humidity are conducive to the growth of fungi associated with the sorghum grain mold (SGM) disease complex (Figure 1). Due to the sustained humidity in existing production environments (Western and Southern Africa, Southern Asia, South America, and Eastern North America), grain mold unfortunately persists throughout the growing season and post-harvest through the off-season if warm and wet storage conditions occur (Tola and Kebede, 2016). Storage conditions conducive to post-harvest mold generally occur when high-moisture grain is not properly dried prior to storage, resulting in storage of wet grain and generally higher levels of non-grain plant residue. Storage facilities that lack proper airflow capacity consequently encourage growth of fungal pathogens such as *Aspergillus* spp. from infected grain and residue (Kange et al., 2015). While SGM remains a threat to yield, it is not consistently the top yield-limiting disease in sorghum as it has been called on occasion in the past. The greater concern in regard to SGM is that the associated mycotoxin contamination is indisputably one of the top international threats to sorghum grain quality and safety. Critical reviews on SGM were published by Castor and Frederiksen (1980), Williams and Rao (1981); Forbes et al. (1992), Bandyopadhyay et al. (2000), and Waniska et al. (2001). However, it has been two decades since an updated review of grain mold literature has been provided, and much progress has been made toward understanding the SGM disease complex from physiological, genetic, and biochemical perspectives.

The SGM pathosystem is exceedingly dynamic, as grain mold is innately multifarious, consisting of a multitude of fungi demonstrating various trophic lifestyles: necrotrophic, saprophytic, and hemibiotrophic. The taxonomic diversity of the SGM complex most commonly encompasses but is not limited to *Fusarium* spp., *Aspergillus* spp., *Curvularia* spp., *Colletotrichum* spp. and, *Alternaria* spp. (Forbes, 1986; Forbes et al., 1992; Bandyopadhyay et al., 2000; Little et al., 2011; Cuevas and Prom, 2020). The various fungal constituents that comprise the hierarchical makeup of the SGM complex drastically fluctuates throughout sorghum growth, harvest, and storage. The physiological changes that occur as the host transitions from vegetative to reproductive stages as well as the shift from the growing season into post-harvest storage environments affect the fungal constitution of the SGM complex (Forbes et al., 1992;

Bandyopadhyay et al., 2000). The ascomycete *F. verticillioides*, previously known as *F. moniliforme* (Seifert et al., 2003), is the single most predominant species during the growing season (Williams and Rao, 1981; Cuevas et al., 2019a). *Aspergillus* however, is regarded as a top threat to the storage of grain, flourishing in poor storage conditions with high moisture or the presence of insects (Amaike and Keller, 2011; Gemedé and University, 2016).

During early host anthesis (i.e., flowering), fungal constituents of the grain mold complex can infect and colonize spikelet tissue prior to grain development (Forbes et al., 1992), resulting in “blasting” of the grain and poor seed set. Post-anthesis, grain mold can externally colonize intact grain outside of the pericarp; however, true damage arises from internal colonization that can hinder grain fill and be detrimental to grain quality. Internal infection of grain is generally a consequence of opportunistic fungal infection at sites of deterioration due to weathering of developing or mature grain. Consequences of internal infection of grain are: (1) digestion of the starches and proteins within the endosperm; (2) overall softening and decay of the seed; and (3) the excretion of toxic secondary metabolites called mycotoxins into the caryopsis, with the latter having one of the most detrimental effects to overall quality and safety (Rodriguez-Herrera et al., 2000). In brief, exposure to dietary mycotoxins has been associated with health adversities such as cancer and cirrhosis (Wild and Gong, 2010; Ostry et al., 2017), immunological disorders (Corrier, 1991), and impaired child growth indicators (Lombard, 2014). In India, a foodborne diarrheal disease outbreak in the 1990s was linked to the consumption of mycotoxin-contaminated sorghum and maize (*Zea mays*) (Bhat et al., 2008). In addition to human diseases, about 48% of total sorghum grain is used as animal feed (Astoreca et al., 2019). In livestock, economically important pathologies associated with mycotoxins, such as Turkey X disease, and other exposure-associated nutritional deficiencies, are pervasive in global animal production systems (Bhat et al., 2010).

Worldwide, mycotoxin exposure through food is widespread (Astoreca et al., 2019) and the amount of international crops affected by mycotoxins was previously estimated near 25% (Park et al., 1999). Park et al. (1999) cited (Boutrif and Canet, 1998) as supporting evidence – even though no mention of this 25% estimate is present in the 1998 publication. This issue was recognized and addressed by Eskola et al. (2019), concluding that the original estimate of 25% could be based on EU legislation and Codex Alimentarius limits. The authors argue this estimate could largely understate the percentage of total food crops that could be infected with mycotoxins at a detectable level, which is argued by (Eskola et al., 2019) to be upward of 60–80%. Estimates for economic losses as a result of SGM-related damages have been previously estimated at US \$130 million globally (Das et al., 2020), however, the United States Department of Agriculture (USDA) Grain Fungal Diseases and Mycotoxin Reference states economic losses due to related mycotoxins can be difficult to accurately assess (Schmale and Munkvold, 2009; USDA-ARS, 2018a,b).

The USDA estimates that 300–400 mycotoxins have been identified to date (USDA, 2016), with a portion of these being



FIGURE 1 | Sorghum grain mold affecting the panicles of tannin and non-tannin cultivars.

notable as they are produced by various pathogens implicated in the SGM complex and adversely affect marketable grain quality to an extent to which they are a severe threat to international food and feed safety (Abrunhosa et al., 2016); aflatoxins, ochratoxins, zearalenone (ZEN), fumonisins, deoxynivalenol (DON) and patulin are commonly listed among this subset (Bandyopadhyay et al., 2000; Bhat et al., 2010; Omotayo et al., 2019). Sorghum is the fifth most-produced cereal globally and a staple food in many parts of the world, such as Africa and Asia. Consequently, sorghum grain is responsible for a considerable portion of the food-related mycotoxin exposure as described above. However, unlike other cereals, sorghum does not yet have legislation regulating the maximum mycotoxin concentration in commercial grain (Astoreca et al., 2019).

SGM-ASSOCIATED FUNGI AND THEIR MYCOTOXIN BIOSYNTHETIC POTENTIAL

The pathogenic relationships between SGM-associated fungi and their sorghum host have been difficult to characterize due to the dynamic spatiotemporal development and multi-organismal nature of the disease complex. SGM has generally been regarded as a single entity, yet the complex does not have a single causal agent. Rather, SGM is a syndrome attributable to a diverse assemblage of fungal taxa (Mpofu and McLaren, 2014), the composition of which can be highly variable across

regions (Table 1). Additionally, the fungal makeup of the complex is heavily influenced by fluctuating dominance among species in the hierarchy throughout the lifespan of the host (Hareesh et al., 2000), which ultimately has a direct effect on the potential for specific production of various mycotoxins. Thus, both pathogen-host and pathogen-pathogen interactions are important determinants of SGM disease outcomes, further convoluting prospects for developing stable host resistance.

The diverse fungal species associated with SGM each take on unique physical and biochemical relationships with their host, involving both pathogenic and saprophytic lifestyles. Given the breadth and complexity of these relationships, host molecular defense mechanisms and the underlying genetics have been difficult to characterize. The need to breakdown the multi-species complex and scrutinize its various components has been expressed in previous literature (Das et al., 2011; Mpofu and McLaren, 2014), and numerous efforts have been made in recent years (Cuevas et al., 2019a; Nida et al., 2019; Prom et al., 2020a). However, substantial gaps remain in understanding these relationships, reflected by lack of deployed SGM resistance in sorghum grown in agricultural environments around the world. Additionally, little is known about the ecology and mycotoxin biosynthetic potential of SGM-associated fungi, further limiting the extent to which conclusions can be drawn about the risk of toxin exposure within and across environments. The recent successful efforts in breaking down the components of the disease complex, and understanding pathogenicity on sorghum hosts, have created an opportunity to more critically examine the roles

TABLE 1 | Summary of SGM-associated fungi implicated in biosynthesis of some major mycotoxins.

Mycotoxin	SGM-associated Fungi	Region	References
Aflatoxin	<i>Aspergillus flavus</i> , <i>A. parasiticus</i>	Egypt	Abdel-Sater et al., 2017
		Tunisia	Lahouar et al., 2016
		India	Mukherjee and Lakshminarasimham, 1995
Deoxynivalenol	<i>Fusarium graminearum</i>	Australia	Blaney and Dodman, 1988
		India	Ramakrishna et al., 1989
Fumonisin	<i>Fusarium</i> Section <i>Liseola</i>	United States	Leslie et al., 1996
		Italy	Moretti et al., 1995
		Spain	Visconti and Doko, 1994
		Nigeria	Vismer et al., 2019
		Philippines	Leslie et al., 1996
		Burundi	Munimbazi and Bullerman, 1996
		India	Sharma et al., 2011
		Australia	Nelson et al., 1992
		Argentina	Pena et al., 2019
Moniliformin	<i>Fusarium</i> Section <i>Liseola</i>	United States	Leslie et al., 1996
		South Africa	Leslie et al., 1996
		Philippines	Leslie et al., 1996
Nivalenol	<i>Fusarium graminearum</i>	Australia	Blaney and Dodman, 1988
Ochratoxin A	<i>Aspergillus</i> Section <i>Circumdati</i> , <i>Penicillium</i> spp.	South Africa	Gil-Serna et al., 2011
		Tunisia	Lahouar et al., 2017
Zearalenone	<i>Fusarium graminearum</i> , <i>F. semitectum</i> , <i>F. incarnatum</i>	Australia	Blaney and Dodman, 1988
		India	Gupta, 1998
		Japan	Aoyama et al., 2014
		Tunisia	Lahouar et al., 2017

References and regions correspond to studies confirming mycotoxin biosynthetic ability of sorghum-associated fungal isolates *in vitro*.

of the implicated fungal taxa in the disease complex and their implications for the future of research regarding SGM.

***Fusarium* spp.**

A predominant genus of SGM-associated fungi is *Fusarium*, which contains numerous mycotoxigenic species and is among the most common genera in the disease complex (Williams and Rao, 1981). The most common species associated with sorghum grain mold are *F. verticillioides* (synonymous with *F. moniliforme* in the literature for the purposes of this review), *F. thapsinum*, and *F. proliferatum*, although many others have also been isolated from sorghum grain (Pena et al., 2019). *F. verticillioides* and

related taxa produce the mycotoxin fumonisin, which has been implicated in human and animal diseases such as esophageal cancer, equine leukoencephalomalacia, and impaired growth (Chen et al., 2018). In more temperate growing environments, DON, a mycotoxin produced by *F. graminearum*, is a major concern across susceptible cereal crops (Das et al., 2011; Dweba et al., 2017). Sorghum contamination with DON has been occasionally reported (Das et al., 2011), but *F. graminearum* is demonstrably less virulent in the disease complex than other fusaria, such as *F. thapsinum* (Mpofu and McLaren, 2014).

Mycotoxigenic Characteristics of Sorghum-Associated *Fusaria*

Relatively little is known about the mycotoxin biosynthetic potential of *Fusarium* isolates derived specifically from sorghum. Among isolates of *Fusarium verticillioides* derived from sorghum, much variability has been documented in fumonisin B1 (FB1) biosynthetic potential. Nelson et al. (1991) showed that among 15 sorghum- or millet-derived *F. moniliforme* isolates from sub-Saharan Africa, 6 produced detectable FB1 *in vitro* with a range of 95–2448 µg/g when incubated at 25°C for 31 days (Nelson et al., 1991). In South Africa, by contrast, it has been shown that *F. napiforme* isolates associated with sorghum molds produce very little to no FB1 (Nelson et al., 1992). In Kansas, USA, FB1-producing *Fusarium* isolates produced between 3 and 3148 µg/g FB1 when cultured on cracked maize under the same conditions (Leslie, 1992).

Moniliformin (MON), another mycotoxin produced by *Fusarium* species associated with SGM, can also accumulate in sorghum and be cause for human and animal health concern. In Argentina, sorghum-derived *F. thapsinum* isolates produced 175–2,100 µg/kg MON (mean 579 µg/kg) on sorghum medium *in vitro* when incubated at 28°C for 21 days (Pena et al., 2019). Leslie et al. (1996) found that sorghum-derived *F. verticillioides* isolates from United States, Philippines, and South Africa were capable of producing MON when incubated on ground maize at 25°C for 21 days. While the isolate with the highest biosynthetic potential (10,345 µg/kg) was collected in South Africa, vast ranges in MON production by *F. verticillioides* was observed across the isolates of the various geographical regions (Leslie et al., 1996).

Aside from fumonisins and moniliformin, the *Fusarium* genus is a prolific producer of trichothecene mycotoxins such as deoxynivalenol (DON) and zearalenone (ZEN), and others, which have been documented in sorghum. DON biosynthetic potential of sorghum-derived *Fusarium* isolates has also been studied to some depth in a range of environments. In India, isolates cultured in glucose-yeast extract-peptone broth at 28°C for 14 days exhibited relatively low levels (0.01–0.04 µg/g) of DON biosynthesis (Ramakrishna et al., 1989). By contrast, a *F. graminearum* isolate from Queensland, Australia produced 239 µg/g on ground maize substrate incubated at 28°C for 28 days (Blaney and Dodman, 1988). This divergence in toxigenicity across geographies is consistent with earlier evidence that populations of *F. graminearum* are differential in their abilities to produce toxins and cause disease in their plant hosts (Ward et al., 2008). Aoyama et al. (2015) found that Japanese isolates of

F. semitectum derived from sorghum hosts produce substantial amounts of ZEN on sorghum substrate, ranging from 1,530 to 19,400 µg/g. This is higher than what has been observed among isolates of *F. semitectum* and other *Fusaria* in India, where ZEN production among toxigenic strains ranged from 0.6 to 2.4 µg/g on rice substrate when incubated at 25°C for 20 days (Gupta, 1998).

Fumonisin as a Virulence Factor in Host Diseases

Fusarium verticillioides, a major contributor to SGM disease outcomes globally, exhibits both biotrophic and necrotrophic modes of infection. With respect to its role in the sorghum grain mold disease complex, *F. verticillioides* infection typically occurs via aerial spores that colonize developing panicles during florescence (Williams and Rao, 1981). There has been debate regarding whether fumonisins act as virulence factors in *Fusarium*-related plant diseases (Munkvold, 2003; Dastjerdi and Karlovsky, 2015). While the infection of *F. verticillioides* and bioactivity of FB1 within the host has not been extensively studied in sorghum, the etiology of infection via FB1 and fumonisins as a whole has garnered attention in other crops such as maize. There is some evidence, largely from seedling diseases, that fumonisin-producing isolates are more virulent than isolates that do not produce fumonisins (Gilchrist, 1998; Abbas et al., 2000; Bacon et al., 2008). Abbas et al. (2000) for example, demonstrated that FB1 is active on contact and exhibits limited translocation via the xylem when able to penetrate through a wound, thus playing an active role in necrotrophic infection. However, there has been little convincing evidence that fumonisins play a role in *Fusarium* virulence in grain diseases (grain molds, ear rots, etc.).

Several studies have been unable to demonstrate a relationship between fumonisin production and disease outcomes. Dastjerdi and Karlovsky (2015) reported from stalk rot of maize that fumonisin-producing and non-producing strains of *F. verticillioides* were shown to colonize stalks at an equal rate and caused similar levels of ear rot. Additionally, both fumonisin producing and non-producing strains were able to create similar levels of maize ear rot. Igarashi et al. (2013) found in Arabidopsis that the successful suppression of FB1-related host cell death caused by host recognition of FB1 is independent of any phytohormonal signaling response, and is actually a pathogen-associated molecular pattern-triggered immunity. In maize treated with *F. verticillioides*, Campos-Bermudez et al. (2013) found that the transcriptional and subsequent metabolic responses of the susceptible variety were far more accentuated than the resistant, suggesting that metabolite synthesis was actually supporting *F. verticillioides* growth and alluded to a more constitutive form of defense in resistant genotypes (Campos-Bermudez et al., 2013). This statement aligns with current literature as FB1 elicits a salicylic acid (SA)-based response in the host – a response usually reserved for responding to biotrophic invasion (Glazebrook, 2005; Verwaaijen et al., 2019). This SA response results in induced apoptosis via hypersensitive reaction (Torre-Hernandez et al., 2010), a successful mode of defense against a biotroph, but counterproductive in the case of necrotrophic and hemibiotrophic infection, ultimately creating

dead host tissue that is optimal for necrotic *F. verticillioides* colonization. These studies illustrate the potential relationship (if any) of fumonisin production to *F. verticillioides* infection (Munkvold, 2003).

Aspergillus spp.

The *Aspergillus* genus contains among the most prolific producers of the mycotoxins aflatoxin (AF) and ochratoxin. The *Aspergillus* species common in the SGM disease complex include soil-borne saprophytes that can opportunistically infect the developing caryopsis following events such as insect damage (Pfliegler et al., 2020). The *Aspergilli* commonly found in sorghum grain molds are: *A. flavus*, *A. niger*, *A. parasiticus*, *A. fumigatus*, and *A. glaucus* (Little et al., 2011; Gemed and University, 2016). *A. flavus* and *A. parasiticus* are the most notable producers of aflatoxins, and while *A. fumigatus* produces the immunosuppressive mycotoxin gliotoxin to increase virulence to plants, animals and humans, it is not known to produce aflatoxin (Kamei and Watanabe, 2005). The risks associated with aflatoxin exposure are most pronounced in tropical and subtropical environments, where *Aspergillus* spp. proliferate abundantly, and the risk is less in regions with both cooler and drier off seasons (Cotty et al., 1994; Amaike and Keller, 2011). Additionally, growing regions with the inability to utilize modern grain storing technologies, which boast increased aeration and drying of stored grain, are at risk of enhanced fungal growth during storage.

The consumption of aflatoxin-contaminated grain can lead to aflatoxicoses (Marin et al., 2013; Mahato et al., 2019), which are some of the most widespread mold-associated diseases globally (Latgei, 1999). Aflatoxin exposure can result in cancer, lack of immune suppression, liver damage, and mortality (Bennett and Klich, 2003; Sarma et al., 2017). From a public health view and in regard to the SGM complex, *A. flavus* and its close relative within the *Flavi* section, *A. parasiticus*, are of the highest concern within the *Aspergillus* genus (Sarma et al., 2017). Compared to *A. parasiticus*, *A. flavus* has a wide host range, and is much more prevalent in the SGM complex across locations (Amaike and Keller, 2011; Little et al., 2011; Gemed and University, 2016; del Palacio and Pan, 2020). Out of the more than 20 known aflatoxins there are four major aflatoxins which are a focus of studies for their abundance in foods and toxicity: B1, B2, G1, G2, and M1 (Iqbal et al., 2015; Kumar et al., 2017). AFB1 is the most toxic of these four, being a potent genotoxic and hepatotoxic agent as well as being classified as a group 2A carcinogen by the International Agency for Research on Cancer (IARC) (Khoury et al., 2011). The *A. flavus* S and L strains are prolific producers of aflatoxins B1 and B2, and in addition, the L strains produce aflatoxins G1 and G2 (Amaike and Keller, 2011).

Under *Aspergillus* infection, sorghum has been shown to produce varying levels of antifungal proteins including sormatin, glucanases, and chitinases (Seetharaman et al., 1997; Ratnavathi and Sashidhar, 2004). Ratnavathi and Sashidhar (2004) showed chitinase as a response mechanism to *Aspergillus* from relatively strong positive correlations between chitinase activity and aflatoxin levels for both white-pericarp ($r^2 = 0.482$) and red-pericarp ($r^2 = 0.600$) sorghums. The authors added that the lower

rate of chitinase activity in red to white-seeded sorghums is likely due to red-seeded sorghums reliance on a constitutively higher presence of polyphenols. Chitinase and glucanase production in an effort to fragment a pathogen cell wall is a common defense mechanism in host plants (de Ferreira and da Monteiro, 2017), and while further studies in sorghum are limited, chitinase production has been shown to be an effective defense response to *Aspergillus* in maize and peanut (Fountain et al., 2015; Hawkins et al., 2015).

In addition to their roles as pathogens in the SGM disease complex, *Aspergilli* also constitute a serious post-harvest spoilage threat for sorghum and other crops when grain is stored under sub-optimal conditions conducive for fungal growth and mycotoxin accumulation. Post-harvest genetic resistance is especially important in ensuring the prevention of aflatoxin infection throughout storage, as *Aspergillus* remains a top threat to storage safety. Sorghum is a staple crop in many developing countries which lack proper storage facilities capable of limiting contamination, and post-harvest genetic resistance is an important defense measure which prevents further expenses such as fumigation from occurring (Gemede and University, 2016; Meseka et al., 2018; Soni et al., 2020).

Compared to the *Aspergilli* associated with maize and groundnuts, the crops most vulnerable to aflatoxin accumulation in many environments, there is relatively little understanding of the mycotoxigenic characteristics of sorghum-associated *Aspergillus* strains. In parts of the world where sorghum is a major staple, notably India and sub-Saharan Africa, some evidence has accrued enabling characterization of aflatoxin biosynthetic potential of sorghum-derived isolates. In Egypt, Abdel-Sater et al. (2017) found that local sorghum-derived isolates of *Aspergillus flavus* produced an average of 254 µg/kg AFB1 when incubated on potato dextrose agar at 28°C for 10 days. Tunisian sorghum-derived isolates, on the other hand, produced just 1.15 µg/kg AFB1 on whole sorghum grains incubated at 37°C for 7 days (Lahouar et al., 2016). This discrepancy is consistent with earlier evidence that the thermal optimum for aflatoxin production is ~30°C (Mousa et al., 2011), but could also be indicative of differential mycotoxigenicity of fungal isolates even from the same geographic region.

In India, *Aspergillus flavus* isolates from sorghum grain produced on average 2567 µg/kg AFB1 in yeast extract broth when incubated at 27°C for 7 days (Usha et al., 1994). Other studies in India also report aflatoxin biosynthetic potential in this ballpark in culture media (Somashekar et al., 2004; Reddy et al., 2011). Incubation on sorghum grain as a substrate has yielded much lower levels of aflatoxin production by sorghum-derived *Aspergillus* isolates in India (Mukherjee and Lakshminarasimham, 1995). This is illustrative of the importance of substrate characteristics in aflatoxigenesis; Winn and Lane (1978), for example, observed marked differences in aflatoxin deposition across sorghum substrates in different forms (whole, cracked, and ground) (Winn and Lane, 1978).

***Curvularia* spp.**

Another ascomycete genus commonly associated with SGM is *Curvularia*, particularly *C. lunata*, which has been shown

to play an increased role in the disease complex in drier environments and to have greater negative impact on germination in early season (Prom et al., 2016). Similarly, to *F. verticillioides*, *Curvularia* is typically associated with early-stage infection of the developing sorghum caryopsis and is known to elicit defense response genes in the host plant (Little and Magill, 2003). There is some evidence that *Curvularia* may be more competitive in the disease complex than *Fusarium* spp. in moist environments and that infection by this fungus substantially reduces seed viability even without presenting severe symptoms of infection (Prom et al., 2003). While certainly regarded as an important disease agent in SGM, there is no confirmed evidence of mycotoxin deposition by *Curvularia* in sorghum and it is unlikely that this genus plays an active role in sorghum mycotoxin contamination.

***Colletotrichum* spp.**

The *Colletotrichum* genus, also the causal agent of another economically important sorghum disease, anthracnose, is known to play a role in the SGM disease complex. The predominant species implicated in both pathosystems is *C. graminicola*, which is abundant in warm, humid sorghum production contexts and has high levels of genetic diversity (Cuevas et al., 2016). While *Colletotrichum* spp. can occasionally predominate over other taxa in the SGM disease complex (Bandyopadhyay et al., 2000), the relative importance of this species is likely highly dependent on climatic conditions in the growing environment. In Texas, a low-humidity environment, for example, there is evidence that *Colletotrichum* spp. is a far less prominent member of the fungal assemblage than *Fusarium*, *Curvularia*, or *Alternaria* taxa (Prom et al., 2011). Like *Curvularia* spp., *Colletotrichum* is likely not implicated directly in mycotoxigenesis in the SGM disease complex. However, it remains unknown the extent to which atoxigenic taxa involved in SGM symptom manifestation modulate or enable toxin deposition by toxigenic fungi via disruption of physical, biochemical, or immune defenses of the host – this should be the subject of future investigation into the host-pathogen and pathogen-pathogen interactions within the disease complex.

Alternaria* spp. and *Epicoccum sorghinum

While not known to infect early flowering tissue of sorghum, *A. alternata* and *A. solani* are common contributors to SGM disease symptoms and have also been implicated in mycotoxin contamination (Bandyopadhyay et al., 2000). In certain environments, such as Texas and Turkey, this genus can be the most prominent member of the SGM fungal assemblage (Prom et al., 2020a). *Alternaria* is a diverse genus of ascomycete fungi, which is a prolific mycotoxin producer and is implicated in a range of plant diseases, notably in high-value crops (Logrieco et al., 2009). Like *Fusarium* spp., the mode of infection of *Alternaria* spp. in the SGM disease complex is largely via airborne spores, which are present in

the growing environment throughout caryopsis development (Bandyopadhyay et al., 1991). Sorghum-associated *Alternaria* can be prolific producers of key mycotoxins such as tenuazonic acid, alternariol, alternariol methyl, altenuene, and altertoxin. The aforementioned *Alternaria* mycotoxins can be mutagenic and cytotoxic and, while less toxic than aflatoxins or fumonisins, are believed to have synergistic effects (Patriarca et al., 2007). At present, tenuazonic acid contamination in sorghum- or millet-based infant foods is the only governmentally regulated *Alternaria* toxin (Tralamazza et al., 2018).

Sorghum contamination with *Alternaria* mycotoxins has been observed in many production contexts globally. Despite the known role of this genus in the disease complex, relatively little attention has been given to the toxigenicity of *Alternaria* associated with SGM, compared to higher-profile toxins in public health dialog. In India, isolates of sorghum-derived *Alternaria tenuissima* produced 32.6, 22.1, 9.2, 2.7, and 1.8 $\mu\text{g/g}$ of these toxins, respectively, on rice medium incubated at 25°C for 28 days (Bilgrami et al., 1995). *Epicoccum sorghinum* (formerly *Phoma sorghinum*) is another important contributor to SGM symptoms in some production systems and has also been implicated in the production of tenuazonic acid. A recent study from Brazil documented substantial tenuazonic acid production by sorghum-derived isolates between 0.0986 and 148 $\mu\text{g/g}$ incubated on rice substrate at 25°C for 21 days (Oliveira et al., 2019).

MECHANISMS OF GRAIN MOLD RESISTANCE IN SORGHUM

Previous publications dating from the 1980s to the early 2000s have described important sources of host resistance to sorghum grain mold, recognized and selected upon through visual features such as panicle compactness (Williams and Rao, 1981; Brown et al., 2006), grain pigmentation representative of tannin content, presence of a pigmented testa, grain hardness, endosperm texture, pericarp thickness, and both glume coverage and tenacity (Williams and Rao, 1981; Menkir et al., 1996; Brown et al., 2006; Sharma et al., 2010). A study by Sharma et al. (2010) using the sorghum mini core collection (Upadhyaya et al., 2009) found two major phenotypes to contribute to SGM resistance in significant ways: panicle compactness and grain pigmentation. Panicle structures exhibiting compactness were found to significantly increase grain mold severity ($r = 0.47$) and grain pigmentation to decrease severity ($r = -0.45$). It should be noted that the relationship between grain pigmentation and SGM is difficult to determine using visual methods due to potential bias, as it is inherently easier to observe mold incidence on white pericarp grain than pigmented grain.

Open panicle structures common in the guinea race of sorghum allow increased resistance to mold colonization by maximizing airflow and minimizing moisture within the panicle. Additionally, guinea sorghum exhibits large glumes with high grain coverage, a trait shown to limit grain mold infestation (Audilakshmi et al., 1999). However,

guinea sorghum commonly demonstrates poor yield as a result of less grain per branch than other sorghum races and exhibit undesirable agronomic traits. Consequently, most commercial varieties consist of semi-open to semi-compact panicle architecture rather than open. This trend demonstrates the limitations of relying on open-panicle structure for SGM resistance.

While panicle architecture is a compilation of types of inflorescence branching, grain pigmentation is a culmination of a diverse and abundant set of biochemical compounds housed within the pericarp, testa, and endosperm (Waniska, 2000). Presence of a dominant spreader gene is necessary for the spread of pigmentation from the testa into the pericarp (Xiong et al., 2019b). Glume pigmentation is also influenced by biochemical compounds similar to the grain. A large portion the biochemical compounds that influence grain and glume color consists of the secondary metabolites that fall into the chemical class of phenolics (Shen et al., 2018), many of which have demonstrated biochemical host resistances that are active against SGM infection. Ubiquitous throughout plants, phenolics of sorghum are responsible for reacting to environmental cues and responding to stress (Cheynier, 2012; Razzaq et al., 2019). Total phenolics in sorghum grain have previously been broken down into phenolic acids, condensed tannins, and flavonoids, all of which have been implicated into SGM resistance at some level (Shen et al., 2018; Xiong et al., 2019b).

Phenolic compounds underlying pigmentation in sorghum demonstrate a range of antioxidant potential that provides an effective source of phytochemical resistance to diseases such as SGM through scavenging of reactive oxygen species generated during fungal infection by both the host and pathogen (Das and Roychoudhury, 2014). While constitutive maintenance of phenolic compounds is widely present in sorghum grain, synthesis of critical phenolic compounds with increased specificity to biotic resistance is largely induced in response to pathogen detection (Liu et al., 2010). Due to this responsiveness of inducible phytochemical production, total phenolic content in sorghum grain unchallenged by pathogens does not effectively predict biotic resistance (Dicko et al., 2005; Atanasova-Penichon et al., 2016).

The presence of a pigmented testa contributes a large amount of condensed tannins and a considerable portion of phenolic content to the biochemical profile of sorghum grain (Menkir et al., 1996). However, the use of condensed tannins within a pigmented testa as a primary source of SGM resistance is limited as tannins lower feed acceptance (Rooney and Murty, 1982) and hinder protein digestion in both animals and humans (Wu et al., 2012), as well as reduce the efficiency of grain starch to ethanol conversion (Zhao et al., 2009). Forbes et al. (1992) stated the need to understand resistance mechanisms of varieties without a pigmented testa (no to low condensed tannins), and since the publishing of this review, progress has been made in elucidating the grain mold – sorghum pathosystem and understanding the host phenolic profile far beyond tannin content. Important sources of both inducible and constitutive phytochemical resistances have been further

characterized; largely enhanced by an improved understanding of the underlying metabolic pathways and processes in sorghum.

BIOACTIVE METABOLITES AND PATHWAYS INVOLVED IN HOST RESPONSE TO GRAIN MOLD INFECTION

Improved knowledge of the transcriptional responses and metabolic architecture of not only sorghum, but crops such as wheat (*Triticum aestivum*) and maize (that encounter similar head mold diseases), has guided the characterization of cereal grains response to grain mold-specific pathogens to limit host stress and mitigate mycotoxins production. Increased characterization of the head mold pathosystems has provided a framework to further understand host phytochemical resistances of sorghum that are derived from metabolic pathways. The following sections seek to highlight specific host chemical mechanisms that are active throughout SGM disease development to provide a general understanding of these interactions to assist in the further research and development of sorghum cultivars with effective phytochemical grain mold resistances while maintaining and even improving important agronomic traits.

In sorghum, grain pigmentation is underlined by a diverse metabolic profile, and has been used to visibly select resistance-related phytochemical traits. Many of these chemical compounds underlying grain pigmentations are phenolic compounds, which represent a large group of secondary metabolites present across the plant kingdom that are synthesized in response to biotic and abiotic stress (Bhattacharya et al., 2010). Various phenolic compounds that play roles in SGM resistance are individual products and intermediaries of the phenylpropanoid biosynthesis pathway. Induced biosynthesis of these phenolics as a defense mechanism occurs predominantly in resistant sorghum genotypes (Katilé et al., 2010; Prom et al., 2020a).

The Phenylpropanoid Biosynthesis Pathway

The phenylpropanoid pathway begins with the enzymatic conversion of phenylalanine by phenylalanine ammonia lyase (PAL) (Tohge et al., 2013), which then undergoes additional enzymatic conversions resulting in naringenin chalcone, the intermediate that chalcone isomerase (CHI) ultimately converts to the flavanone naringenin and marks the beginning of the flavonoid biosynthesis pathway (Figure 2). Many of the phenolic acids contributing to SGM disease resistance are derivatives of phenylalanine, and while part of the phenylpropanoid pathway, are not part of the downstream flavonoid pathway.

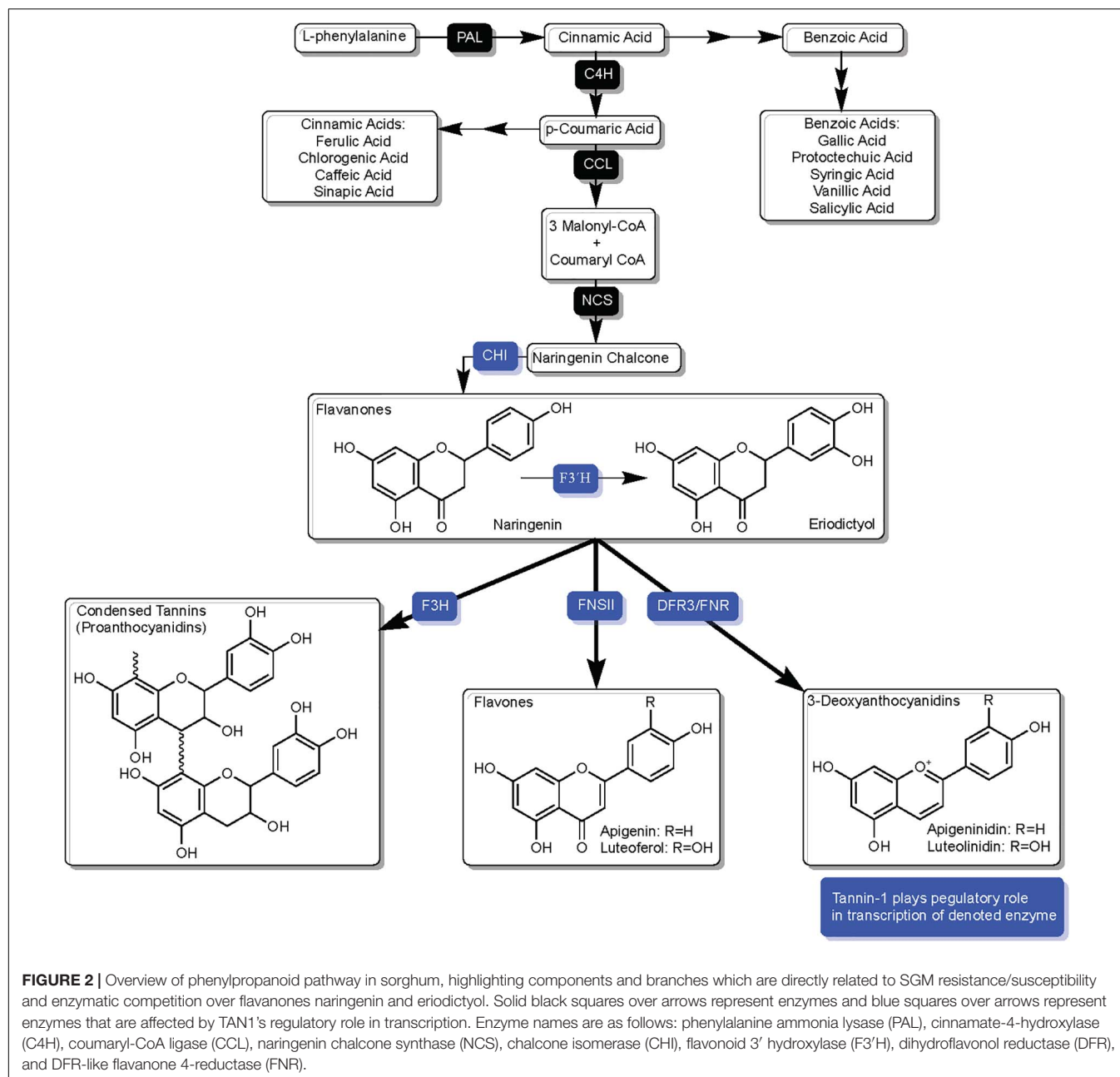
Phenolic acids represent the largest and simplest group of non-flavonoid phenolics in sorghum grain. Having been subject to an array of anti-mycotoxin research, phenolic acids have been shown to both inhibit and activate mycotoxin production (Atanasova-Penichon et al., 2016; Stuper-Szablewska and Perkowski, 2017). The main phenolic acids connected to SGM resistance reported

in sorghum consist of the (1) cinnamic acids: ferulic, chlorogenic, caffeic, *p*-coumaric, and sinapic acids (2) benzoic acids: gallic, protocatechuic, vanillic, and syringic acids (Xiong et al., 2019b). In cereals, cinnamic acids such as ferulic and *p*-coumaric acid have been shown to have inhibitory effects for *Fusarium* growth and mycotoxin production (Ferruz et al., 2016b), and additionally shown to directly inhibit *F. verticillioides* and *F. proliferatum* mycelial proliferation and fumonisin production on maize-based media (Ferrochio et al., 2013). Studies have shown caffeic acid and vanillic acid to drastically reduce FB1 production and mycelial growth of *Fusarium* (Beekrum et al., 2003; Schöneberg et al., 2018). In contrast to cinnamic acids, benzoic acids with the exception of syringic acid have generally been shown to have activating effects, even providing slight stimulation to mycelial growth (Boutigny et al., 2009; Atanasova-Penichon et al., 2016; Schöneberg et al., 2018).

Phenolic acids exist in both bound and free forms. Bound forms represent 70–95% of phenolic acid in sorghum grain (Xiong et al., 2019b) and present a multifaceted grain mold resistance source. Alongside of contributing to antioxidant potential and limiting mycotoxins on a cellular level, bound forms of phenolic acids strengthen grain hardness (Chiremba et al., 2012). Alongside ferulic acid, *p*-coumaric acid has also been shown to positively correlate with sorghum grain hardness (Chiremba et al., 2012). Grain hardness contributes to grain mold resistance by reducing weathering of the pericarp and ultimately limits the ability of grain mold pathogens to degrade the endosperm. This connection of phenolic acids to grain hardness suggests an interesting contribution to SGM-resistance (Jambunathan et al., 1992; Waniska et al., 2001).

Compared to bound phenolic acids, free phenolic acids represent a smaller portion of total phenolic acids within sorghum, maize and wheat. Additionally, biosynthesis levels undergo more drastic fluctuations than their bound counterparts (Atanasova-Penichon et al., 2016). In many cereals, free chlorogenic acid is the most common free phenolic acid. Gauthier et al. (2016) demonstrated in maize that chlorogenic acid is a valuable source of resistance to *F. graminearum* proliferation and limiter of type B trichothecene mycotoxin production. While chlorogenic acid exhibits antifungal properties of its own, the authors showed that *F. graminearum* transforms host chlorogenic acid into caffeic acid, which exhibits antifungal properties of increased potency over the former. The authors concluded this phenomenon regarding chlorogenic acid exhibits “pro-drug” qualities, demonstrating its highest toxicity levels when degraded into caffeic acid (Gauthier et al., 2016). Chlorogenic acid is shown to be a major phenolic acid present in sorghum (Pasha et al., 2015), shown to play roles in essential physiological processes such as photosynthesis (Turner et al., 2016); however, the role of chlorogenic acid as a fungal inhibitor in sorghum has not been explored to the extent that it has in maize or wheat.

Both the strain of pathogen and host are critical determinants in the overall efficiencies of phenolic acids as they interact with pathogens in inhibitory, neutral or activating roles (Ferruz et al., 2016a; Gauthier et al., 2016). Research has helped create a solid understanding of the roles phenolic acids play in diseases such as wheat *Fusarium* head blight and maize



ear rot. Even though sorghum boasts a diverse and abundant phenolic profile well suited for further exploration in relation to biotic resistance, phenolic acid activity has not been as extensively studied with sorghum-based substrates using grain mold pathogens and related mycotoxins. Presence of common phenolic acids across crops and studies showing functional properties of phenolic acids on mycotoxins demonstrates a potential for increased contribution to grain mold resistance in sorghum. To understand the nuances of phenolic acid based-sorghum grain mold resistances, interactions of phenolic acids with grain mold pathogens must undergo further research using sorghum-based media or *in vivo* studies to understand these relationships *in planta*.

The Flavonoid Biosynthesis Pathway: Naringenin as a Precursor Influenced by Host Sensitivity

Naringenin marks the beginning of the flavonoid biosynthesis pathway, acting as a substrate that is subject to competition from an array of enzymes: flavonoid 3' hydroxylase (F3'H), flavanone-3-hydroxylase (F3H) as well as dihydroflavonol 4-reductase (DFR) and the DFR-like enzyme flavanone 4-reductase (FNR) (Xiong et al., 2019a; **Figure 2**). The ultimate fate of naringenin is heavily influenced by genotype by environment interactions, with studies such as (Taleon et al., 2012) demonstrating 48% of total variation in naringenin abundance being a consequence of

environmental variation, and separate studies having reported major influence on naringenin by biotic factors such as fungal ingress (Boddu et al., 2005). The fate of naringenin is also tissue-dependent, as downstream products such as 3-deoxyanthocyanidins (3-DAs) are maintained constitutively in unchallenged grain while synthesis is only induced in leaf tissue upon fungal ingress (Boddu et al., 2005; Ibraheem et al., 2015; Kawahigashi et al., 2016).

Sorghum demonstrates a high level of host sensitivity to fungal ingress, as spatial accuracy in the induction of phytoalexins critical to SGM resistance occur through site-specific synthesis at place of infection. Naringenin is a precursor to many of these phytoalexins and marks an important junction in the flavonoid pathway. A resistant response to fungal ingress is dependent on the availability of naringenin to enzymes that synthesize SGM resistance-related compounds. As described in detail in later sections, synthesis of 3-DA phytoalexins is an inducible resistance-related host response to SGM, and a lack of activity by F3H can be a consequence of inducible synthesis of 3-DAs resulting from increased channeling of DFR3/FNR on naringenin (Figure 2; Liu et al., 2010). As will be explored throughout this text, the availability of naringenin at optimal time and point of infection is crucial to host response. At its core, a large portion of host biochemical responses to grain mold may in many ways a consequence of the management of the enzymatic competition over naringenin.

Enzymatic Conversion of Naringenin by F3H – The Synthesis and Properties of Condensed Tannins (Proanthocyanidins)

Condensed tannins (proanthocyanidins) are well understood for their roles in increased antioxidant capabilities and significant correlations with SGM resistance (Harris and Burns, 1973; Forbes, 1986; Melake-Berhan et al., 1996; Dicko et al., 2005; Cuevas et al., 2019b). The presence of condensed tannins throughout SGM resistant germplasm is interesting, as there is little knowledge to suggest condensed tannins demonstrate direct toxicity to fungal pathogens as do products of other branches of the flavonoid biosynthesis pathway such as 3-DAs and flavones (Melake-Berhan et al., 1996; Du et al., 2009; Nida et al., 2021). Regardless, the use of condensed tannins is limited in sorghum germplasms for their reduction in protein digestibility of animals. However, the toxicity of condensed tannins to humans has been misunderstood, and have been shown to provide antioxidants, fiber and reduce obesity as a food source (Dykes et al., 2005). Sorghum tannin types have been divided amongst three categories: type I) no tannins; type II) tannins in pigmented testa; type III) tannins in pigmented testa and pericarp (presence of spreader gene) (Waniska et al., 2001).

SbF3H1 of sorghum codes the flavanone-3-hydroxylase (SbF3H) enzyme that is capable of converting naringenin to condensed tannins (Figure 3). As another competitive enzyme for naringenin, host management of SbF3H activity could play an important role in managing enzymatic competition for naringenin (Mizuno et al., 2014). SbF3H converts flavanones

to dihydroflavonols from which SbDFR1 synthesizes flavan-3,4-diols (leucoanthocyanidins). Leucoanthocyanidins represents a junction in this branch of flavonoid synthesis from which anthocyanidins or flavan-3-ols can be synthesized. Condensed tannins are the result of oligomerization and polymerization of flavan-3-ol compounds (He et al., 2008; Wu et al., 2012).

Early understandings of SGM resistances recognized condensed tannin content as a lone dominant influence on SGM resistance (Harris and Burns, 1973). While this outlook was justified at the time, it was chemical studies such as Menkir et al. (1996) that began to expand this perspective to consider a more complete host phenolic profile by showing inconsistencies in condensed tannin content of SGM resistant phenotypes. Recent studies have assisted in the accurate understanding of the contributions of condensed tannin content to SGM resistance as the expansive genetic underpinnings of grain mold resistance continue to be elucidated. Biosynthesis of condensed tannins is controlled by the *Tannin-1 (TAN1)* transcription factor (Wu et al., 2012). *TAN1* demonstrates influence over enzymatic competition for naringenin by regulating expression of enzyme-coding genes in the pericarp responsible for the transcription of *CHI*, *F3H*, *F3'H*, *DFR* and *ANS* (Wu et al., 2012; Mizuno et al., 2014), and consequently influencing the synthesis of 3-DAs, flavones, and anthocyanins in addition to condensed tannins (Liu et al., 2010).

Wu et al. (2012) found the *TAN1* allele was present throughout 78% of a diverse collection of sorghum germplasm consisting of 161 accessions. Similar to the 161 accessions used by Wu et al. (2012), Cuevas et al. (2019a) found that a functional *TAN1* is present in 79% of the sorghum association panel (SAP) (Casa et al., 2008), an interesting finding considering that a majority of the accessions within the SAP were also found to be highly susceptible to SGM. The study by Cuevas et al. (2019a) demonstrated that genotypes harboring a non-functional *TAN1* allele (*tan1-a*) were more susceptible to grain mold on average compared to accessions with the functional allele. Interestingly, the Genome-Wide Association Studies (GWAS) of the SAP did not associate *TAN1* with grain mold resistance, illustrating the complexity of the SGM resistance phenotype that supports the author's conclusion that resistance is a culmination of many additional factors in addition to the concentration of tannins (Cuevas et al., 2019a).

Another recent SGM association study employing 635 Ethiopian genotypes by Nida et al. (2021) identified a collection of protein-coding loci, and consistently detected SNPs within the *TAN1* coding region. The authors suggested *TAN1*'s role in the flavonoid pathway affecting biosynthesis of 3-DAs and tannins could contribute to *TAN1*'s role in grain mold resistance. While widespread presence of *TAN1* across mapping populations is a factor of both aforementioned studies, a key difference is the differing levels of resistance between the two populations, with the former (SAP) demonstrating increased susceptibility overall than the latter (Ethiopian). This notable difference in SGM resistance between mapping populations may be a consequence of varying levels of *TAN1*'s role in transcriptional regulation of the enzyme-coding genes in the flavonoid pathway, and could be factoring into why *TAN1* was not associated with

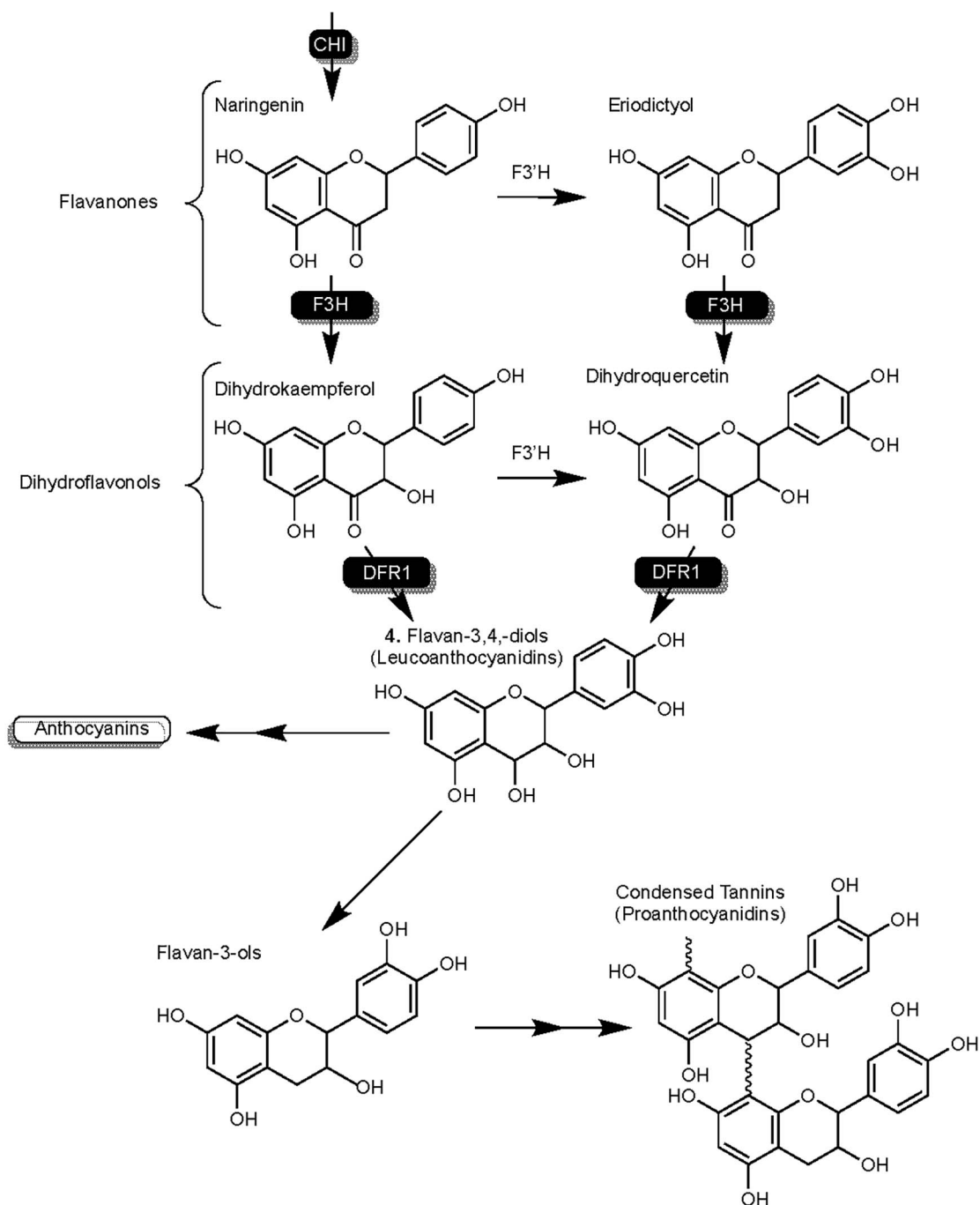


FIGURE 3 | Detailed overview of the flavonoid biosynthesis branch responsible for the synthesis of condensed tannins (proanthocyanidins) in sorghum. Highlights F3H activity on flavanones naringenin and eriodictyol. Enzymes are as follows: chalcone isomerase (CHI), flavanone-3-hydroxylase (F3H), and dihydroflavonol reductase (DFR).

SGM resistance in the SAP, but consistently identified in the Ethiopian panel.

Additional GWAS such as those by Prom et al. (2020a) and Ahn et al. (2019) further highlight the genetic complexities of SGM resistance, as many identified genes show minor rather than major effects, such as a variety of zinc finger proteins, resistance

genes (R-genes), and the underlying genetics associated with systemic acquired resistance (SAR) mechanisms (Ahn et al., 2019; Prom et al., 2020a). While condensed tannin content remains an important component of host resistance, it only represents a portion of the phenolic profile that contributes to a complex host resistance mechanism. Many additional components of host

resistance such as sorghum SAR mechanisms, R-genes and the phenolic profile in its entirety should be explored in order to provide a more complete understanding of mold resistance in sorghum germplasms (Cuevas et al., 2019a). Additionally, at the core of the grain mold resistance phenotype in sorghum is host management of enzymatic competition over naringenin and the ability to ensure naringenin availability to the correct branch of the flavonoid pathway at time of grain mold infection.

Enzymatic Conversion of Naringenin by SBD3R3/FNR – The Synthesis and Properties of Flavan-4-OLS, 3-Deoxyanthocyanidins, and Phlobaphenes

Flavan-4-ols

Flavan-4-ols have been subject to many early studies regarding grain mold resistance, and repeatedly shown to be a critical component of sorghum defense mechanisms active against SGM (Jambunathan and Kherdekar, 1991; Melake-Berhan et al., 1996). Studies have shown resistant genotypes respond to pathogen infection by producing flavan-4-ols, and demonstrated their importance as a precursor and active compound (Jambunathan and Kherdekar, 1991; Poloni et al., 2014). The dihydroflavonol 4-reductase enzyme SBD3R3 and DFR-like flavonoid 3'-hydroxylases (SbFNR) are responsible for the conversion of naringenin into flavan-4-ol (Kawahigashi et al., 2016; Mizuno et al., 2016) (Figure 4). Resistant and susceptible genotypes have been shown to maintain flavan-4-ols at similar concentrations in early sorghum development (Jambunathan and Kherdekar, 1991). Three weeks post-anthesis however, flavan-4-ol content in susceptible genotypes decreased over three times more than in resistant lines (Melake-Berhan et al., 1996). While flavan-4-ol concentrations have been shown to be present in the caryopsis throughout grain mold resistant genotypes, flavan-4-ols do not directly inhibit the growth of fungal pathogens (Schutt and Netzly, 1991), suggesting a role as a precursor or in signaling mechanisms.

3-Deoxyanthocyanidins

3-DAs of sorghum are composed of luteolinidin, apigenidin and their methoxylated derivatives (Xiong et al., 2019a). These anthocyanins demonstrate high levels of fungal toxicity and have been tied to fungal resistance mechanisms (Nicholson et al., 1987; Schutt and Netzly, 1991). Early studies such as the one by Menkir et al. (1996) found the phytoalexin apigenidin to contribute to grain mold resistance and found no relationship between resistance with luteolinidin and apigenidin combined. Interestingly, notable studies had previously determined luteolinidin, when compared to apigenidin, exhibits higher levels of fungal toxicity to *H. maydis* and *C. graminicola*, accumulates at slower rates following infection (Nicholson et al., 1987), and is more consistently present throughout anthracnose resistant varieties (Snyder and Nicholson, 1990). While luteolinidin accumulates at slower rates than apigenidin, literature suggest resistant varieties exhibit increased reliance on this extra enzymatic step involving F3'H to produce luteolinidin, possibly

to take advantage of the higher fungal toxicity over apigenidin (Nicholson et al., 1987; Mizuno et al., 2014). Melake-Berhan et al. (1996) were able to demonstrate high correlations of apigenidin and luteolinidin with each other as well as with flavan-4-ols; however, the study was unable to demonstrate a major contribution of luteolinidin and apigenidin to grain mold resistance, even though both compounds demonstrate effective fungal toxicity.

3-DAs have been found to be present in maize and flowers of the Gesneriaceae and Bignoniaceae families, but pathogen-induced rapid production of 3-DAs as phytoalexins is unique to sorghum (Liu et al., 2010; Xiong et al., 2019a). In sorghum, 3-DAs are produced in infected cells as inclusion bodies (Snyder and Nicholson, 1990), and migrate to the site of infection (Liu et al., 2010). This production and migration of phytoalexin 3-DAs is primarily a pathogen-induced mechanism, and undetectable in tissues of unchallenged hosts. The enzymes SBD3R3 and SbFNR are responsible for the conversion of naringenin into flavan-4-ol precursors, and an unknown ANS enzyme converts flavan-4-ols to 3-DAs (Figure 4; Xiong et al., 2019a).

Much of the research to understand the transcriptional controls and site-specific production of 3-DAs in response to fungal ingress have been performed in relation to anthracnose resistance using *Colletotrichum sublineolum* (Ibraheem et al., 2010; Kawahigashi et al., 2016). Consequently, much of the research that has highlighted these phytoalexins for site-specificity in sorghum analyzed 3-DA biosynthesis in leaf and mesocotyl tissue. Sorghum fungal diseases such as anthracnose predominantly initiate in leaf tissue, while the most damaging aspects of grain mold occur in developing reproductive tissue (i.e., grain). Leaf tissue production of 3-DAs is primarily an induced mechanism and only measurable when challenged by fungal pathogens (Kawahigashi et al., 2016). In contrast, SGM resistant genotypes maintain constitutive, detectable levels of 3-DAs within grain that is unchallenged by SGM (Awika et al., 2004; Boddu et al., 2005).

Two overlying concepts in the understandings of 3-DA activity in the SGM disease complex outline two shortcomings in current understandings: (1) Antifungal properties and pathogen-related induction of 3-DAs suggests a considerable role in grain mold resistance, but evidence of significant contributions to phenotype are inconsistent; and (2) the analysis of inducible 3-DA activity in SGM specific-interactions and host tissues (developing reproductive tissue, grain, and glume) compared to other fungal pathogens of sorghum has been lacking. Recently however, with the addition of association and expression studies, Nida et al. (2019) successfully identified and connected genes responsible for 3-DA synthesis to SGM resistance and identified induced expression of flavonoid biosynthesis-related structural genes in the grain and glume. The authors performed GWAS for grain mold scores across a large set of 2,010 accessions, with 1,940 Ethiopian landrace accessions, 1,550 of which were from the Ethiopian Biodiversity Institute. This GWAS successfully identified the previously mapped locus that contains the MYB R2R3 transcription factor gene *YELLOW SEED1* (Y1) (Nida et al., 2019), that Boddu et al. (2005) had implicated to play a direct role in synthesis of 3-DAs and phlobaphenes in the

speculate as to the ability of phlobaphenes to inactivate fungal proteins by the formation of irreversible complexes as part of host detoxification mechanisms. Additionally, the enzymatic machinery which synthesizes phlobaphenes from flavan-4-ols has not yet been elucidated. While showing promise in resistance to *Fusarium* spp. of ear rot in maize, phlobaphenes are mainly regarded as a contributor to grain pigmentation in sorghum, and have not been scrutinized for their contribution to other phenotypes such as biotic resistance.

Pericarp Color 1 (P1) of maize, an ortholog to sorghum *Y1*, has been shown to be functionally similar in the maize flavonoid biosynthesis pathway, with a confirmed role in the accumulation of phlobaphenes (Ibraheem et al., 2015; Landoni et al., 2020). *Y1* of sorghum, as described earlier for its role in 3-DA biosynthesis, is primarily known in sorghum for control of pericarp pigmentation (Brenton et al., 2016). In addition to pigmentation, results from Boddu et al. (2005) demonstrate that a functional *Y1* is also required for biosynthesis of phlobaphenes. Additionally, in comparison to maize, sorghum phlobaphenes are present in both vegetative and reproductive tissues; whereas, maize phlobaphenes are generally relegated to reproductive tissues (Boddu et al., 2006; Ibraheem et al., 2015). While both *Y1* and a high presence of flavan-4-ols are present across much of the SGM-resistant germplasm (Menkir et al., 1996; Nida et al., 2019), much is to be determined as to the competition for flavan-4-ols as a substrate. Much progress has been made in understanding the presence of 3-DAs across SGM-resistant sorghum germplasm (Dicko et al., 2005; Nithya et al., 2013); however, the alternate fate of flavan-4-ols and the presence of phlobaphenes across SGM-resistant germplasm has yet to be characterized to the same degree.

Enzymatic Conversion of Naringenin by SbFNSII – The Synthesis and Properties of Flavones

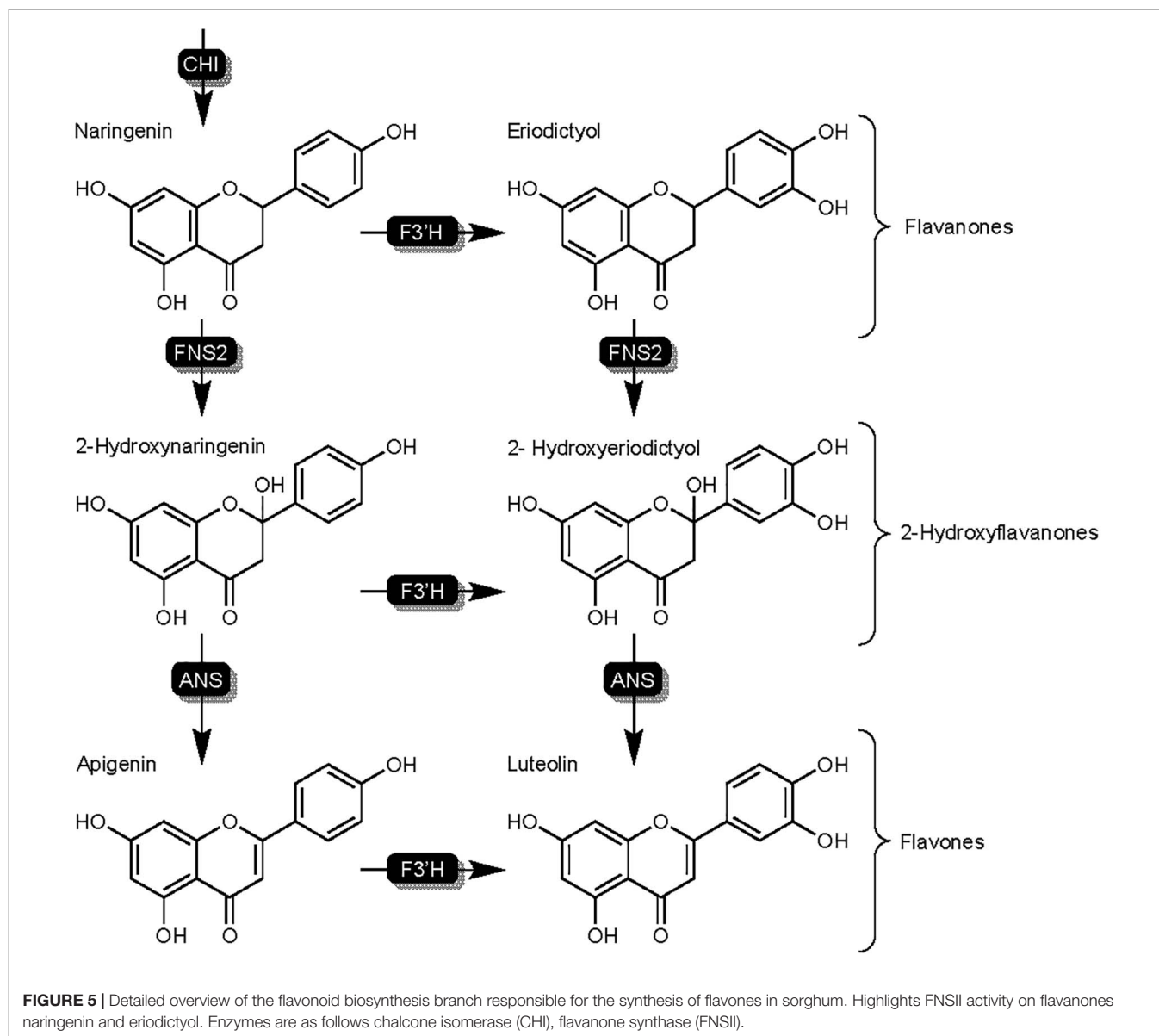
Flavone synthase II (FNSII) enzymes are responsible for the synthesis of the flavones apigenin and luteolin (Figure 5), and like F3'H, are part of the cytochrome P450 superfamily (Yonekura-Sakakibara et al., 2019). Flavones in sorghum represent a separate branch of the flavonoid biosynthesis pathway that begins with the hydroxylation of naringenin or eriodictiol by SbFNSII resulting in the formation of 2-hydroxyflavone intermediates (Du et al., 2009). Subsequent conversion of 2-hydroxyflavones by an unknown ANS results in the formation of the flavones apigenin and luteolin (Figure 5). Flavones are one of the largest subgroups of flavonoids and are well understood for their ability to affect color by complexing with anthocyanins as well as antioxidant potential and nutrient value amongst other benefits (Martens and Mithöfer, 2005). Additionally, flavones such as luteolin and apigenin have been shown to demonstrate levels of fungal toxicity of their own, with increased fungal toxicity and higher presence in resistant germplasm demonstrated by the former (Du et al., 2009; Aboody and Mickymaray, 2020). Research in sorghum has suggested flavones are a phytoalexin due to a combination of the aforementioned fungal toxicity and pathogen-related induction (Du et al., 2009). Functional characterization of flavone synthase

genes in other crops however, has reported an interesting mechanism in balancing the competition between FNSII and other enzymes and even pathogen-manipulation of flavone production potentially being a component of host susceptibility (Ferreira et al., 2015).

Darker pericarp color and purple plant sorghum lines have been generally shown to be more resistant to disease (Melake-Berhan et al., 1996; Menkir et al., 1996; Kawahigashi et al., 2016). Kawahigashi et al. (2016) reported that during tissue wounding, red pericarp, purple plant sorghum types favor conversion of naringenin and eriodictiol by SbFNR (3-DA synthesis), while seemingly suppressing conversion by SbFNSII - suggesting increased reliance on flavan-4-ols and 3-DA phytoalexins pathways over flavones (Dykes et al., 2005; Kawahigashi et al., 2016). On the contrary, tan plants favor conversion of naringenin and eriodictiol by SbFNSII while seemingly blocking the production of flavan-4-ols, resulting in leaf and glume tissue showing higher levels of apigenin and luteolin over 3-DAs (Kawahigashi et al., 2016; Mizuno et al., 2016). This phenomenon could potentially be a factor contributing to the higher antioxidant capabilities and disease resistance of purple plant/dark pericarp sorghum genotypes over tan plants.

The findings described above from Kawahigashi et al. (2016) suggest potential mechanisms of sorghum in managing enzymatic activity on naringenin to balance compound synthesis between separate flavonoid pathway branches. Research in maize and Arabidopsis has begun to show this as an important trait of SA-mediated resistance, a phytohormone commonly associated with SAR responses. Flavone synthase enzymes responsible for flavone synthesis such as FNSII are widely present throughout the flavonoid biosynthesis pathways of plants. FNSI and FNSII enzymes of maize are functionally similar to SbFNSII, as both play roles in the synthesis of flavones from naringenin and both have been shown to play an important role in the SA – flavone cross talk affecting maize salicylic acid-mediated resistance (Ferreira et al., 2015; Righini et al., 2019). It is important to note that FNSI distinctly differs from FNSII in that it is part of the 2-oxoglutarate-dependent dioxygenase family and shows higher sequence identity to F3H than to FNSII (Du et al., 2009; Yonekura-Sakakibara et al., 2019). While maize has been shown to rely on both ZmFNSI and ZmFNSII, sorghum has thus far been shown to be reliant on SbFNSII for flavone synthesis (Righini et al., 2019). An important dissimilarity in the roles of FNSI and FNSII exists: FNSI converts flavanones directly to flavones, while most FNSII enzymes convert flavanones to 2-hydroxyflavanones prior to conversion to flavones by an unknown protein (Figure 5; Du et al., 2009; Kawahigashi et al., 2016).

ZmFNSI-1 is commonly coexpressed with maize R2R3-MYB transcription regulator *P1* (Ferreira et al., 2015), an ortholog of the critical SGM resistance gene *Y1* (Boddu et al., 2005; Nida et al., 2019). Transcriptional studies show *ZmFNSI-1* is highly active in maize pericarps and silk - tissues commonly under pressure from ear rot-related pathogens (Righini et al., 2019). In maize and Arabidopsis, a SA – flavone cross talk that fine-tunes homeostasis between the



phytohormones and flavones has been well described. Studies have characterized an *Arabidopsis* SA 5-hydroxylase DMR6 (Downy mildew resistance), a recently reported homolog of ZmFNSI-1 (Zhang et al., 2017b). In *Arabidopsis*, the ZmFNSI-1 homolog DMR6 was shown to be responsible for controlling levels of salicylic acid (SA) and naringenin through use of either compound as a substrate. However, studies show that DMR6 prefers SA as a substrate over naringenin, showing that increased availability of SA over naringenin to DMR6/ZmFNSI-1 will consequently result in higher availability of naringenin. The authors conclude that while enhanced SA levels are responsible for pathogen resistance, lowered levels of apigenin and luteolin are likely not directly connected to resistance. Rather, the resistant mechanism is a result of the preoccupation of ZmFNSI-1/DMR6 with the hydroxylation of SA rather than naringenin,

increasing the availability of naringenin as a substrate to other resistance-related enzymes to increase phytoalexin production (Zhang et al., 2017b).

Metabolic studies of sorghum under *Burkholderia andropogonis* (leaf blight) infection point toward induction of flavone synthase as a defense response (Mareya et al., 2019), while studies in *Arabidopsis* and maize have suggested the potential of flavones in active immune suppression through metabolic manipulation by the pathogen resulting in host susceptibility (Ferreyra et al., 2015). Studies regarding pigmented sorghums suggest a role in limiting SbFNSII activity to support resistance related enzymes such as SbFNR (Dykes et al., 2005; Kawahigashi et al., 2016). Additionally, recent SGM association studies identified BTB/POZ domains within SAR response pathways that involve phytohormone induction (Cuevas et al., 2019a). Rather than SA-induced resistance

as described above, Liu et al. (2010) found that expression of *SbDFR3* and 3-DA pathways were induced by methyl jasmonate treatment of sorghum roots and antagonized by SA. The enzymatic competition for naringenin in sorghum and coexpression of *Y1* ortholog *P1* with *ZmFNSI-1* in maize suggest flavones may be carefully managed in SGM interactions; however, host mechanisms must be further characterized in reproductive tissue using grain mold pathogens to elucidate this relationship.

With confirmation of the role of *Y1* in SGM resistance, the combined increased expression of *SbDFR3* and elucidation of implicated SAR mechanisms identified in SGM resistant sorghum suggests there is potential for heightened enzymatic competition for naringenin and a phytohormone-mediated mechanism for management of enzymatic activity on naringenin (Nida et al., 2019; Prom et al., 2020a). The presence of flavones in SGM resistant varieties has been demonstrated, but further functional characterization of *SbFNSII* in SGM pathosystems is required to understand to which extent that host reliance on flavone synthesis branches of the flavonoid pathway over *SbDFR3/SbFNR* and *F3H*-mediated branches contributes to a SGM resistant phenotype.

Enzymatic Conversion of Naringenin by F3'H – The Synthesis and Properties of Eriodyctiol and 3' Hydroxylated Derivatives

Part of the cytochrome P450 superfamily, flavonoid 3'-hydroxylase enzyme (F3'H) plays a potentially important role in the SGM response of sorghum as it is responsible for the conversion of naringenin to the flavanone intermediate eriodyctiol (Figure 2; Poloni et al., 2014; Yonekura-Sakakibara et al., 2019). Through hydroxylation of the 3' position of the B-ring of naringenin, F3'H synthesizes eriodyctiol, the precursor to luteolinidin and luteolin (Figures 4, 5; Mizuno et al., 2016). Luteolinidin has been shown to be present at a higher rate than apigenidin throughout red pericarp sorghum, a pigment trait that correlates with SGM resistance (Dykes et al., 2005). Additionally, luteolinidin exhibited increased fungal toxicity over apigenidin (Nicholson et al., 1987). Luteolin has also been described to exhibit higher presence in disease resistant sorghum lines as well (Du et al., 2009). This phenomenon is suggestive that the extra enzymatic step in F3'H hydroxylation to produce eriodyctiol as an intermediate may play a role in enhancing resistance mechanisms to SGM pathogens.

MOVING FORWARD WITH ELUCIDATING SGM HOST-PATHOGENS INTERACTIONS AND DEVELOPING SGM RESISTANT CULTIVARS

Research has shown grain mold resistance of sorghum to be a vastly complex and highly quantitative trait (Cuevas et al., 2019a). Recent efforts have begun to isolate the fungal

pathogens of the grain mold complex and differentiate host responses. A wealth of knowledge has been provided by recent association studies, but more research is needed to characterize these host-pathogen interactions in both tissue-specific and pathogen-inducible contexts. As more QTL are implicated into SGM resistance, understanding host manipulation of metabolic pathways and pathogen biosynthetic mycotoxin potential is vital to fully elucidating the SGM disease complex and host response mechanisms.

As fungal assemblages change throughout regions, the plasticity of grain mold resistance in current germplasm justifies increased attention. Many studies have been performed over the years screening germplasm in a variety of regions under the stress of different fungal assemblages of the SGM complex (Table 2). As warmer air and widespread humidity becomes more present in the face of climate change, the importance of acknowledging different fungal assemblages in regions with newly developing grain mold pressure will become increasingly important. Much of sorghum grain production relies on hybrid varieties, and considerable positive heterosis for grain mold resistance has been demonstrated by breeding programs (Kumar et al., 2008; Hundekar et al., 2015). Hundekar et al. (2015) reported a better parent heterosis of SGM resistance of 147.22%, but also reported a negative heterotic effect as far as -17.32%. Studies have reported success in developing grain mold resistant hybrid varieties through the use of resistant varieties as the female parent (Thakur et al., 2008), citing the many minor QTL responsible for SGM resistance as a potential factor. While the value of heterosis in SGM resistance has been demonstrated, little is known as to how this phenomenon affects the many biochemical mechanisms described throughout this text. More research is required to better understand the heterotic contributions to the metabolic makeup of sorghum in relation to host resistance.

Many of the underlying transcription factors related to SGM resistance like *Y1* and *TAN1* have shown widespread transcriptional regulation of a multitude of biochemical compounds. Orthologous genes such as *P1* of maize demonstrate both similar and dissimilar functions, as divergence of orthologous genes can be significant even in species such as maize and sorghum with a close evolutionary relationship (Zhang et al., 2017a). Consequently, while functional characterization and enhanced understanding of head molds in other crops such as maize and wheat demonstrate a framework for hypotheses development in sorghum, extrapolation must be approached with caution when researching the transcriptional and metabolic responses of SGM.

Many additional topics related to the understanding of the SGM disease complex suffer from a lack of general knowledge in sorghum. Included within this group is mycotoxin detoxification *in planta*, knowledge of the contribution of heterosis to host plant resistance, changes in host response through various physiological growth stages, relationships between the individual pathogens in the fungal makeup of SGM, and the effects of fungal endophytes on host resistance. Fungal endophytes have demonstrated the ability to reprogram the host transcriptome and metabolome to favor secondary metabolism in ryegrass (Dupont et al., 2015). While similar

TABLE 2 | Table highlighting projects that have screened grain mold resistance in sorghum germplasm at different locations worldwide, under differing fungal assemblages (if listed), and report resistant varieties.

References	Population	Number of lines	Area of growth	Pathogens detected/used if listed
Bandyopadhyay et al., 1988	World Collection	7,132	ICRISAT, Patancheru, India	N/A
Audilakshmi et al., 1999	Selected lines from NRCS, AICSIP, and ICRISAT	22	Directorate of Oil Seeds Research, Rajendranagar, Hyderabad, India	N/A
Reddy, 2000	ICRISAT - hybrid and parental lines	N/A	ICRISAT, Patancheru, India	N/A
Prom et al., 2003	Selected lines	8	Texas A&M Agricultural Research Farm, College Station, Texas	<i>F. thapsinum</i> , <i>C. lunata</i>
Prom, 2004	Selected lines	8	Texas A&M Agricultural Research Farm, College Station, Texas	<i>Fusarium thapsinum</i> , <i>F. semitectum</i> , <i>Curvularia lunata</i> , <i>Alternaria</i> spp.
Kumar et al., 2008	ICRISAT - hybrid and parental lines	25	ICRISAT, Patancheru, India	N/A
Thakur et al., 2008	Reevaluation of resistant lines from Bandyopadhyay et al. (1988)	156	ICRISAT, Patancheru, India	N/A
Katilé et al., 2010	Selected lines	12	Texas A&M Agricultural Research Farm, College Station, TX, United States and Greenhouse	<i>Fusarium thapsinum</i> , <i>Curvularia lunata</i> , <i>Alternaria</i> spp., <i>Fusarium semitectum</i> , <i>Aspergillus</i> spp., <i>Rhizopus</i> spp.
Sharma et al., 2010	Sorghum Mini-Core	140	ICRISAT, Patancheru, India	N/A
Prom et al., 2014	US Comercial Hybrids w/breeding line checks	28	Agronomic Research Stations in Bambey and Nioro, Sengal, Africa	N/A
Hundekar et al., 2015	AICSIP - hybrids and parental Lines	44	Sorghum Improvement Project, University of Agricultural Sciences, Dharwad	N/A
Cuevas et al., 2016	Exotic germplasm accessions from Burkina Faso and South Africa	80	Isabela, Puerto Rico	<i>Fusarium thapsinum</i> , <i>Fusarium semitectum</i> , and <i>Curvularia lunata</i>
Prom et al., 2016	Sorghum Conversion Program lines and inbred checks	65	Texas A&M AgriLife Research Farm, Burleson County, Texas	<i>F. thapsinum</i> and <i>C. lunata</i>
Prom et al., 2017	US Breeding lines and hybrids, Senegalese lines	23	Agronomic Research Stations in Bambey, Senegal, West Africa	<i>F. thapsinum</i>
Cuevas et al., 2019a	Sorghum Association Panel	331	USDA- ARS Tropical Agriculture Research Station at Isabela and Mayaguez, PR	<i>Curvularia lunata</i> , <i>Fusarium semitectum</i> , and <i>F. thapsinum</i>
Diatla et al., 2019	Sengalese Institute of Agricultural Research - hybrid and parental lines	10	Darou-Pakathiar, Sinthiou-Damba, Guirigara, Sinthiou-Damba, Center and South-East of Senegal, West Africa	N/A
Nida et al., 2019	Ethiopian sorghum landrace collection, breeding lines, and released varieties	1412, 1414	Bako Agricultural Research Center, Ethiopia.	<i>F. proliferatum</i> , <i>F. graminearum</i> , <i>F. thapsinum</i> , <i>F. verticillioides</i> , <i>F. oxysporum</i> , <i>Alternaria</i>
Nida et al., 2021	Ethiopian sorghum landrace collection	635	Bako and Jimma Agricultural Research Centers, Western Ethiopia	<i>Phoma</i> , <i>Curvularia</i> , <i>Fusarium</i> , <i>Cladosporium</i> , <i>Bipolaris</i> , <i>Alternaria</i> , <i>Colletotrichum</i> and <i>Rhizopus</i>
Prom et al., 2020a	Sorghum Association Panel	377	Texas A&M AgriLife Research Farm, Burleson County, Texas	<i>A. alternata</i> , <i>Fusarium thapsinum</i> , and <i>Curvularia lunata</i>
Prom et al., 2020b	Selected lines from USDA-GRIN	47	Texas AgriLife Research Farm, Burleson County, Texas	<i>A. alternata</i> alone or a mixture of <i>A. alternata</i> , <i>F. thapsinum</i> and <i>C. lunata</i>

USDA – ARS, United States Department of Agriculture – Agricultural Research Service; ICRISAT, Crops Research Institute for Semi-Arid Tropics; AICSIP, All India Co-ordinated Sorghum Improvement Project; NRCS, National Research Center for Sorghum (India).

systemic effects have not been shown in sorghum, a trend such as this would result in increased host favoritism of the many components of the flavonoid pathway that are related to SGM resistance, hypothetically resulting in increased SGM resistance at a potential cost of yield due to a change in photosynthate partitioning. While few endophyte-related host systemic changes have been shown in relation to sorghum host resistances, localized biocontrol of SGM pathogens by fungal endophytes has been demonstrated (Rajini et al., 2020). Rajini et al. (2020) showed biocontrol effects were

demonstrated by 26 endophytic isolates on common SGM complex pathogens: *Fusarium thapsinum*, *Epicoccum sorghinum*, *Alternaria alternata*, and *Curvularia lunata*. This study by Rajini et al. (2020) demonstrates the fungal endophyte potential for localized control, however, characterization of host systemic changes caused by the presence of fungal endophytes requires increased research to understand the potential effects on sorghum biotic resistance.

SGM is poised to increase in relevance as regions boasting warm temperatures and humidity increases due to climate

change. For this reason, increased efforts are required to understand the plasticity of the grain mold complex between environments, how this affects the fungal hierarchy, and ultimately how these changes affect SGM pathogenicity on sorghum as a host. While the use of individual pathogens to isolate host response is critical to research, understanding the relationships within the disease complex is critical to understanding SGM and finding sources of resistance adapted to SGM complexes native to specific regions in international germplasm. Maintaining grain quality as a food and feed and increasing SGM resistance are both one in the same. Achieving this objective is achievable largely in part to sorghums abundant phenolic profile. However, due to the intricate and changing makeup of the SGM disease complex, continued research ventures into SGM show both notable obstacles and great potential for sorghum improvement in increased resilience, quality, and yield.

REFERENCES

- Abbas, H. K., Smeda, R. J., Gerwick, B. C., and Shier, W. T. (2000). Fumonisin B1 from the fungus *Fusarium moniliforme* causes contact toxicity in plants: evidence from studies with biosynthetically labeled toxin. *J Nat Toxins* 9, 85–100.
- Abdel-Sater, M., Abdel-Hafez, S., Hussein, N., and al-amery, E. (2017). Fungi associated with maize and sorghum grains and their potential for amylase and aflatoxins Production. *Egypt J. Botany* 57, 119–137. doi: 10.21608/ejbo.2017.296.1008
- Aboody, M. S. A., and Mickymaray, S. (2020). Anti-Fungal Efficacy and Mechanisms of Flavonoids. *Antibiotics* 9, 45. doi: 10.3390/antibiotics9020045
- Abrunhosa, L., Morales, H., Soares, C., Calado, T., Vila-Chã, A. S., Pereira, M., et al. (2016). A Review of Mycotoxins in Food and Feed Products in Portugal and Estimation of Probable Daily Intakes. *Crit Rev Food Sci* 56, 249–265. doi: 10.1080/10408398.2012.720619
- Agrios, G. N. (2005). "Parasitism and Disease Development," in *Plant Pathology*, ed. A. Abbas 78.
- Ahn, E., Hu, Z., Perumal, R., Prom, L. K., Odvody, G., Upadhyaya, H. D., et al. (2019). Genome wide association analysis of sorghum mini core lines regarding anthracnose, downy mildew, and head smut. *Plos One* 14:e0216671. doi: 10.1371/journal.pone.0216671
- Amai, K., and Keller, N. P. (2011). *Aspergillus flavus*. *Annu Rev Phytopathol* 49, 107–133. doi: 10.1146/annurev-phyto-072910-095221
- Aoyama, K., Ishikuro, E., Noriduki, H., and Ichino, M. (2015). Formation Ratios of Zearalanone, Zearalenols, and Zearalanols versus Zearalenone during Incubation of *Fusarium semitectum* on Sorghum and Ratios in Naturally Contaminated Sorghum. *Food Hyg Saf Sci Shokuhin Eiseigaku Zasshi* 56, 247–251. doi: 10.3358/shokueishi.56.247
- Astoreca, A. L., Emateguy, L. G., and Alconada, T. M. (2019). Fungal contamination and mycotoxins associated with sorghum crop: its relevance today. *Eur J Plant Pathol* 155, 381–392. doi: 10.1007/s10658-019-01797-w
- Atanasova-Penichon, V., Barreau, C., and Richard-Forget, F. (2016). Antioxidant Secondary Metabolites in Cereals: Potential Involvement in Resistance to *Fusarium* and Mycotoxin Accumulation. *Front Microbiol* 7:566. doi: 10.3389/fmicb.2016.00566
- Audilakshmi, S., Stenhouse, J. W., Reddy, T. P., and Prasad, M. V. R. (1999). Grain mould resistance and associated characters of sorghum genotypes. *Euphytica* 107, 91–103. doi: 10.1023/a:1026410913896
- Awika, J. M., Rooney, L. W., and Waniska, R. D. (2004). Properties of 3-Deoxyanthocyanins from Sorghum. *J Agr Food Chem* 52, 4388–4394. doi: 10.1021/jf049653f
- Bacon, C. W., Glenn, A. E., and Yates, I. E. (2008). *Fusarium Verticillioidea*: Managing The Endophytic Association With Maize For Reduced Fumonisin Accumulation. *Toxin Rev* 27, 411–446. doi: 10.1080/15569540802497889
- Balint-Kurti, P., Shew, H. D., and Cowger, C. (2016). "Host Resistance," in *Plant Pathology Concepts and Laboratory Exercises*, eds B. H. Ownley and R. N. Trigiano (London).
- Bandyopadhyay, R., Butler, D. R., Chandrashekar, A., Reddy, R. K., and Navi, S. S. (2000). "Biology, Epidemiology, and Management of Sorghum Grain Mold," in *Proceeding of the Technical and institutional options for sorghum grain mold management: proceedings of an international consultation*.
- Bandyopadhyay, R., Mughugh, L. K., and Rao, K. E. P. (1988). Sources of Resistance to Sorghum Grain Molds. *Plant Disease* 72, 504–508.
- Bandyopadhyay, R., Mughugh, L. K., Satyanarayana, M. V., and Kalisz, M. E. (1991). Occurrence of airborne spores of fungi causing grain mould over a sorghum crop. *Mycol. Res.* 95, 1315–1320. doi: 10.1016/S0953-7562(09)80583-2
- Beekrum, S., Govinden, R., Padayachee, T., and Odhav, B. (2003). Naturally occurring phenols: a detoxification strategy for fumonisin B1. *Food Addit Contam* 20, 490–493. doi: 10.1080/0265203031000098678
- Bennett, J. W., and Klich, M. (2003). Mycotoxins. *Clin Microbiol Rev* 16, 497–516. doi: 10.1128/cmr.16.3.497-516.2003
- Bhat, R., Rai, R. V., and Karim, A. A. (2010). Mycotoxins in Food and Feed: Present Status and Future Concerns. *Compr Rev Food Sci F* 9, 57–81. doi: 10.1111/j.1541-4337.2009.00094.x
- Bhat, R. V., Shetty, P. H., Amruth, R. P., and Sudershan, R. V. (2008). A Foodborne Disease Outbreak Due to the Consumption of Moldy Sorghum and Maize Containing Fumonisin Mycotoxins. *J Toxicol Clin Toxicol* 35, 249–255. doi: 10.3109/15563659709001208
- Bhattacharya, A., Citovsky, V., and Sood, P. (2010). The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. *Mol Plant Pathol* 11, 705–719. doi: 10.1111/j.1364-3703.2010.00625.x
- Bilgrami, K. S., Sinha, K. K., Ansari, A. A., and Rahman, M. F. (1995). Widespread occurrence of toxigenic *Alternaria* in cereals and oilseeds. *Indian Phytopathology* 48, 150–153.
- Blaney, B., and Dodman, R. (1988). Production of the mycotoxins zearalenone, 4-deoxynivalenol and nivalenol by isolates of *Fusarium graminearum* groups 1 and 2 from cereals in Queensland. *Crop Pasture Sci* 39, 21–29. doi: 10.1071/ar9880021
- Boddu, J., Jiang, C., Sangar, V., Olson, T., Peterson, T., and Chopra, S. (2006). Comparative Structural and Functional Characterization of Sorghum and Maize Duplications Containing Orthologous Myb Transcription Regulators of 3-Deoxyflavonoid Biosynthesis. *Plant Mol Biol* 60, 185–199. doi: 10.1007/s11103-005-3568-1

AUTHOR CONTRIBUTIONS

AA and AW drafted the manuscript. RB is the professor advisor who oversaw the creation of the manuscript and provided final edits. All authors contributed to the article and approved the submitted version.

FUNDING

Financial support was awarded to RB by the USDA National Institute of Food and Agriculture Organic Agriculture Research and Extension Initiative (OREI) under grant #2018-51300-28431.

ACKNOWLEDGMENTS

We thank Dr. Rebecca Nelson (Cornell University) for useful discussion and input in the creation of this review.

- Boddu, J., Svabek, C., Ibraheem, F., Jones, A. D., and Chopra, S. (2005). Characterization of a deletion allele of a sorghum Myb gene yellow seed1 showing loss of 3-deoxyflavonoids. *Plant Sci* 169, 542–552. doi: 10.1016/j.plantsci.2005.05.007
- Boutigny, A.-L., Barreau, C., Atanasova-Penichon, V., Verdal-Bonnin, M.-N., Pinson-Gadais, L., and Richard-Forget, F. (2009). Ferulic acid, an efficient inhibitor of type B trichothecene biosynthesis and Tri gene expression in Fusarium liquid cultures. *Mycol Res* 113, 746–753. doi: 10.1016/j.mycres.2009.02.010
- Boutrif, E., and Canet, C. (1998). Mycotoxin prevention and control: FAO programmes. *Revue de Medecine Veterinaire* 149, 681–694.
- Brenton, Z. W., Cooper, E. A., Myers, M. T., Boyles, R. E., Shakoor, N., Zielinski, K. J., et al. (2016). A Genomic Resource for the Development, Improvement, and Exploitation of Sorghum for Bioenergy The designation of a functional model grass species and the. *Genetics* 204, 21–33. doi: 10.1534/genetics.115.183947
- Brown, P. J., Klein, P. E., Bortiri, E., Acharya, C. B., Rooney, W. L., and Kresovich, S. (2006). Inheritance of inflorescence architecture in sorghum. *Theor Appl Genet* 113, 931–942. doi: 10.1007/s00122-006-0352-9
- Campos-Bermudez, V. A., Fauguel, C. M., Tronconi, M. A., Casati, P., Presello, D. A., and Andreo, C. S. (2013). Transcriptional and Metabolic Changes Associated to the Infection by Fusarium verticillioides in Maize Inbreds with Contrasting Ear Rot Resistance. *Plos One* 8:e61580. doi: 10.1371/journal.pone.0061580
- Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., et al. (2008). Community resources and strategies for association mapping in sorghum. *Crop Sci* 48, 30–40. doi: 10.2135/cropsci2007.02.0080
- Castor, L. L., and Frederiksen, R. A. (1980). Fusarium head blight occurrence and effects on sorghum yield and grain characteristics in Texas. *Plant Disease* 64, 1017–1019.
- Chen, C., Riley, R. T., and Wu, F. (2018). Dietary Fumonisin and Growth Impairment in Children and Animals: A Review. *Compr Rev Food Sci F* 17, 1448–1464. doi: 10.1111/1541-4337.12392
- Cheyrier, V. (2012). Phenolic compounds: from plants to foods. *Phytochem Rev* 11, 153–177. doi: 10.1007/s11101-012-9242-8
- Chiremba, C., Taylor, J. R. N., Rooney, L. W., and Beta, T. (2012). Phenolic acid content of sorghum and maize cultivars varying in hardness. *Food Chem* 134, 81–88. doi: 10.1016/j.foodchem.2012.02.067
- Corrier, D. E. (1991). Mycotoxicosis: mechanisms of immunosuppression. *Vet Immunol Immunop* 30, 73–87. doi: 10.1016/0165-2427(91)90010-a
- Cotty, P. J., Bayman, P., Egel, D. S., and Elias, K. S. (1994). "Agriculture, Aflatoxins And Aspergillus," in *The Genus Aspergillus. of European Microbiological Societies Symposium Series*, eds K. A. Powell, A. Renwick, and J. F. Peberdy (Boston, MA: Springer).
- Cuevas, H. E., Fermin-Pérez, R. A., Prom, L. K., Cooper, E. A., Bean, S., and Rooney, W. L. (2019a). Genome-Wide Association Mapping of Grain Mold Resistance in the US Sorghum Association Panel. *Plant Genome* 12, 1–11. doi: 10.3835/plantgenome2018.09.0070
- Cuevas, H. E., and Prom, L. K. (2020). Evaluation of genetic diversity, agronomic traits, and anthracnose resistance in the NPGS Sudan Sorghum Core collection. *Bmc Genomics* 21:88. doi: 10.1186/s12864-020-6489-0
- Cuevas, H. E., Prom, L. K., and Cruet-Burgos, C. M. (2019b). Genome-Wide Association Mapping of Anthracnose (Colletotrichum sublineolum) Resistance in NPGS Ethiopian Sorghum Germplasm. *G3 Genes Genomes Genetics* 9, g3.400350.2019. doi: 10.1534/g3.119.400350
- Cuevas, H. E., Prom, L. K., Isakeit, T., and Radwan, G. (2016). Assessment of sorghum germplasm from Burkina Faso and South Africa to identify new sources of resistance to grain mold and anthracnose. *Crop Prot* 79, 43–50. doi: 10.1016/j.cropro.2015.10.007
- Das, I., Aruna, C., and Tonapi, V. A. (2020). *Sorghum Grain Mold Research*. Telangana: ICAR-Indian Institute of Millets Research.
- Das, I. K., Audilakshmi, S., and Patil, J. V. (2011). Fusarium grain mold: The Major Component of Grain Mold in Sorghum (Sorghum bicolor L. Moench). *European Journal of Plant Science and Biotechnology* *vol and page,
- Das, K., and Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Front. Environ. Sci.* 2:53. doi: 10.3389/fenvs.2014.00053
- Dastjerdi, R., and Karlovsky, P. (2015). Systemic Infection of Maize, Sorghum, Rice, and Beet Seedlings with Fumonisin-Producing and Nonproducing Fusarium verticillioides Strains. *Plant Pathology J* 31, 334–342. doi: 10.5423/ppj.oa.05.2015.0088
- del Palacio, A., and Pan, D. (2020). Occurrence and toxigenic potential of Aspergillus section Flavi on wheat and sorghum silages in Uruguay. *Mycology* 11, 147–157. doi: 10.1080/21501203.2020.1752321
- Diatta, C., Tovignan, T. K., Adoukonou-Sagbadja, H., Aidara, O., Diao, Y., Sarr, M. P., et al. (2019). Development of sorghum hybrids for stable yield and resistance to grain mold for the Center and South-East of Senegal. *Crop Prot* 119, 197–207. doi: 10.1016/j.cropro.2019.02.001
- Dicko, M. H., Gruppen, H., Barro, C., Traore, A. S., van Berkel, W. J. H., and Voragen, A. G. J. (2005). Impact of Phenolic Compounds and Related Enzymes in Sorghum Varieties for Resistance and Susceptibility to Biotic and Abiotic Stresses. *J Chem Ecol* 31, 2671–2688. doi: 10.1007/s10886-005-7619-5
- Du, Y., Chu, H., Wang, M., Chu, I. K., and Lo, C. (2009). Identification of flavone phytoalexins and a pathogen-inducible flavone synthase II gene (SbFNSII) in sorghum. *J Exp Bot* 61, 983–994. doi: 10.1093/jxb/erp364
- Dupont, P.-Y., Eaton, C. J., Wargent, J. J., Fechtner, S., Solomon, P., Schmid, J., et al. (2015). Fungal endophyte infection of ryegrass reprograms host metabolism and alters development. *New Phytologist* 208, 1227–1240. doi: 10.1111/nph.13614
- Dweba, C. C., Figlan, S., Shimelis, H. A., Motaung, T. E., Sydenham, S., Mwadzingeni, L., et al. (2017). Fusarium head blight of wheat: Pathogenesis and control strategies. *Crop Prot* 91, 114–122. doi: 10.1016/j.cropro.2016.10.002
- Dykes, L., Rooney, L. W., Waniska, R. D., and Rooney, W. L. (2005). Phenolic Compounds and Antioxidant Activity of Sorghum Grains of Varying Genotypes. *J Agr Food Chem* 53, 6813–6818. doi: 10.1021/jf050419e
- Eskola, M., Kos, G., Elliott, C. T., Hajšlová, J., Mayar, S., and Krška, R. (2019). Worldwide contamination of food-crops with mycotoxins: Validity of the widely cited 'FAO estimate' of 25%. *Crit Rev Food Sci* 60, 2773–2789. doi: 10.1080/10408398.2019.1658570
- de Ferreira, R. M. S. B. and da Monteiro, S. A. V. (2017). "Plant-Fungal Interactions at the Molecular Level," in *Plant Pathology Concepts and Laboratory Exercises*, eds B. H. Ownley and R. N. Trigiano (Boca Raton, FL: CRC Press), 381–399.
- Ferreira, M. L. F., Emiliani, J., Rodriguez, E. J., Campos-Bermudez, V. A., Grotewold, E., and Casati, P. (2015). The Identification of Maize and Arabidopsis Type I FLAVONE SYNTHASES Links Flavones with Hormones and Biotic Interactions. *Plant Physiol* 169, 1090–1107. doi: 10.1104/pp.15.00515
- Ferrochio, L., Cendoya, E., Farnochi, M. C., Massad, W., and Ramirez, M. L. (2013). Evaluation of ability of ferulic acid to control growth and fumonisin production of Fusarium verticillioides and Fusarium proliferatum on maize based media. *Int J Food Microbiol* 167, 215–220. doi: 10.1016/j.ijfoodmicro.2013.09.005
- Ferruz, E., Atanasova-Pénichon, V., Bonnin-Verdal, M., Marchegay, G., Pinson-Gadais, L., Ducos, C., et al. (2016a). Effects of Phenolic Acids on the Growth and Production of T-2 and HT-2 Toxins by Fusarium langsethiae and F. sporotrichioides. *Molecules* 21, 449. doi: 10.3390/molecules21040449
- Ferruz, E., Loran, S., Herrera, M., Gimenez, I., Bervis, N., Rцена, C., et al. (2016b). Inhibition of Fusarium Growth and Mycotoxin Production in Culture Medium and in Maize Kernels by Natural Phenolic Acids. *J Food Protect* 79, 1753–1758. doi: 10.4315/0362-028x.jfp-15-563
- Forbes, G. (1986). *Characterization of Grain Mold Resistance in Sorghum [Sorghum bicolor (L. Moench)]*. *city pub.
- Forbes, G. A., Bandyopadhyay, R., and Garcia, G. (1992). "A Review of Sorghum Grain Mold," in *Sorghum and millets diseases: a second world review*, eds W. A. J. de Milliano, R. A. Frederiksen, and G. D. Bengston.
- Fountain, J. C., Khera, P., Yang, L., Nayak, S. N., Scully, B. T., Lee, R. D., et al. (2015). Resistance to Aspergillus flavus in maize and peanut: Molecular biology, breeding, environmental stress, and future perspectives. *Crop J* 3, 229–237. doi: 10.1016/j.cj.2015.02.003
- Gauthier, L., Bonnin-Verdal, M.-N., Marchegay, G., Pinson-Gadais, L., Ducos, C., Richard-Forget, F., et al. (2016). Fungal biotransformation of chlorogenic and caffeic acids by Fusarium graminearum: New insights in the contribution of phenolic acids to resistance to deoxynivalenol accumulation in cereals. *Int J Food Microbiol* 221, 61–68. doi: 10.1016/j.ijfoodmicro.2016.01.005
- Gemed, H. F., and University, A. A. (2016). Study on Aspergillus Species and Aflatoxin Levels in Sorghum (Sorghum bicolor L.) Stored for Different Period

- and Storage System in Kewet Districts, Northern Shewa, Ethiopia. *Food Sci Nutrition* 2, 1–8. doi: 10.24966/fsn-1076/100010
- Gilchrist, D. G. (1998). PROGRAMMED CELL DEATH IN PLANT DISEASE: The Purpose and Promise of Cellular Suicide. *Annu Rev Phytopathol* 36, 393–414. doi: 10.1146/annurev.phyto.36.1.393
- Gil-Serna, J., Vázquez, C., Sardiñas, N., González-Jaén, M. T., and Patiño, B. (2011). Revision of ochratoxin A production capacity by the main species of *Aspergillus* section Circumdati. *Aspergillus steynii* revealed as the main risk of OTA contamination. *Food Control* 22, 343–345. doi: 10.1016/j.foodcont.2010.06.018
- Glazebrook, J. (2005). Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annu Rev Phytopathol* 43, 205–227. doi: 10.1146/annurev.phyto.43.040204.135923
- Gupta, R. (1998). Incidence of zearalenone producing *Fusarium* species on sorghum grains. *Indian Phytopathology* *vol page,
- Hareesh, P., Rao, V., Giridhar, P., and Reddy, S. M. (2000). Fungal succession on Sorghum seed harvested at different stages of maturity. *Indian Pathology* *vol page,
- Harris, H. B., and Burns, R. E. (1973). Relationship Between Tannin Content of Sorghum Grain and Preharvest Seed Molding 1. *Agron J* 65, 957–959. doi: 10.2134/agronj1973.00021962006500060033x
- Hawkins, L. K., Mylroie, J. E., Oliveira, D. A., Smith, J. S., Ozkan, S., Windham, G. L., et al. (2015). Characterization of the Maize Chitinase Genes and Their Effect on *Aspergillus flavus* and Aflatoxin Accumulation Resistance. *Plos One* 10:e0126185. doi: 10.1371/journal.pone.0126185
- He, F., Pan, Q.-H., Shi, Y., and Duan, C.-Q. (2008). Biosynthesis and Genetic Regulation of Proanthocyanidins in Plants. *Molecules* 13, 2674–2703. doi: 10.3390/molecules13102674
- Hundekar, R., Kamatar, M. Y., and Brunda, S. M. B. (2015). Heterosis for Yield and Grain Mold Resistance in Rainy Season Sorghum [*Sorghum bicolor* (L.) Moench]. *Environment and Ecology* 34, 1570–1576.
- IARC. (2012). Economics of mycotoxins: evaluating costs to society and cost-effectiveness of interventions. *IARC Sci Publ* 119–129.
- Ibraheem, F., Gaffoor, I., and Chopra, S. (2010). Flavonoid phytoalexin-dependent resistance to anthracnose leaf blight requires a functional yellow seed1 in Sorghum bicolor. *Genetics* 184, 915–926. doi: 10.1534/genetics.109.111831
- Ibraheem, F., Gaffoor, I., Tan, Q., Shyu, C.-R., and Chopra, S. (2015). A Sorghum MYB Transcription Factor Induces 3-Deoxyanthocyanidins and Enhances Resistance against Leaf Blights in Maize. *Molecules* 20, 2388–2404. doi: 10.3390/molecules20022388
- Igarashi, D., Bethke, G., Xu, Y., Tsuda, K., Glazebrook, J., and Katagiri, F. (2013). Pattern-triggered immunity suppresses programmed cell death triggered by fumonisin B1. *Plos One* 8:e60769. doi: 10.1371/journal.pone.0060769
- Iqbal, S. Z., Jinap, S., Pirouz, A. A., and Faizal, A. R. A. (2015). Aflatoxin M1 in milk and dairy products, occurrence and recent challenges: A review. *Trends Food Sci Tech* 46, 110–119. doi: 10.1016/j.tifs.2015.08.005
- Jambunathan, R., and Kherdekar, M. S. (1991). Flavan-4-ol concentration in leaf tissues of grain mold susceptible and resistant sorghum plants at different stages of leaf development. *J Agr Food Chem* 39, 1163–1165. doi: 10.1021/jf00006a035
- Jambunathan, R., Kherdekar, M. S., and Stenhouse, J. W. (1992). Sorghum grain hardness and its relationship to mold susceptibility and mold resistance. *J Agr Food Chem* 40, 1403–1408. doi: 10.1021/jf00020a023
- Kamei, K., and Watanabe, A. (2005). *Aspergillus* mycotoxins and their effect on the host. *Med Mycol* 43, 95–99. doi: 10.1080/13693780500051547
- Kange, A. M., Cheruiyot, E. K., Ogendo, J. O., and Arama, P. F. (2015). Effect of sorghum (*Sorghum bicolor* L. Moench) grain conditions on occurrence of mycotoxin-producing fungi. *Agric Food Secur* 4, 15. doi: 10.1186/s40066-015-0034-4
- Katilé, S. O., Perumal, R., Rooney, W. L., Prom, L. K., and Magill, C. W. (2010). Expression of pathogenesis-related protein PR-10 in sorghum floral tissues in response to inoculation with *Fusarium thapsinum* and *Curvularia lunata*. *Mol Plant Pathol* 11, 93–103. doi: 10.1111/j.1364-3703.2009.00580.x
- Kawahigashi, H., Kasuga, S., Sawada, Y., Yonemaru, J., Ando, T., Kanamori, H., et al. (2016). The Sorghum Gene for Leaf Color Changes upon Wounding (P) Encodes a Flavanone 4-Reductase in the 3-Deoxyanthocyanidin Biosynthesis Pathway. *G3 Genes Genomes Genetics* 6, 1439–1447. doi: 10.1534/g3.115.026104
- Khoury, A. E., Atoui, A., Rizk, T., Lteif, R., Kallassy, M., and Lebrihi, A. (2011). Differentiation between *Aspergillus flavus* and *Aspergillus parasiticus* from Pure Culture and Aflatoxin-Contaminated Grapes Using PCR-RFLP Analysis of aflR-aflJ Intergenic Spacer. *J Food Sci* 76, M247–M253. doi: 10.1111/j.1750-3841.2011.02153.x
- Kumar, A. A., Thakur, R., Reddy, B. V., and Ramaiah, B. (2008). Improved sorghum hybrids with grain mold resistance. *Open Access Journal published by ICRISAT* 6, 1–4.
- Kumar, P., Mahato, D. K., Kamle, M., Mohanta, T. K., and Kang, S. G. (2017). Aflatoxins: A Global Concern for Food Safety, Human Health and Their Management. *Front Microbiol* 7:2170. doi: 10.3389/fmicb.2016.02170
- Lahouar, A., Marin, S., Crespo-Sempere, A., Saïd, S., and Sanchis, V. (2016). Effects of temperature, water activity and incubation time on fungal growth and aflatoxin B1 production by toxigenic *Aspergillus flavus* isolates on sorghum seeds. *Revista Argentina De Microbiol* 48, 78–85. doi: 10.1016/j.ram.2015.10.001
- Lahouar, A., Marin, S., Crespo-Sempere, A., Saïd, S., and Sanchis, V. (2017). Influence of temperature, water activity and incubation time on fungal growth and production of ochratoxin A and zearalenone by toxigenic *Aspergillus tubingensis* and *Fusarium incarnatum* isolates in sorghum seeds. *Int. J. Food Microbiol.* 242, 53–60. doi: 10.1016/j.ijfoodmicro.2016.11.015
- Landoni, M., Puglisi, D., Cassani, E., Borlini, G., Brunoldi, G., Comaschi, C., et al. (2020). Phlobaphenes modify pericarp thickness in maize and accumulation of the fumonisin mycotoxins. *Sci Rep-uk* 10, 1417. doi: 10.1038/s41598-020-58341-8
- Latgei, J.-P. (1999). *Aspergillus fumigatus* and Aspergillosis. *Clin Microbiol Rev* 12, 310–350. doi: 10.1128/cmr.12.2.310
- Leslie, J. F. (1992). Fumonisin B 1 Production by Strains from Different Mating Populations of *Gibberella fujikuroi* (*Fusarium* Section Liseola). *Phytopathology* 82, 341. doi: 10.1094/phyto-82-341
- Leslie, J. F., Marasas, W. F., Shephard, G. S., Sydenham, E. W., Stockenström, S., and Thiel, P. G. (1996). Duckling toxicity and the production of fumonisin and moniliformin by isolates in the A and E mating populations of *Gibberella fujikuroi* (*Fusarium moniliforme*). *Appl Environ Microb* 62, 1182–1187. doi: 10.1128/aem.62.4.1182-1187.1996
- Little, C. R., and Magill, C. W. (2003). Elicitation of defense response genes in sorghum floral tissues infected by *Fusarium thapsinum* and *Curvularia lunata* at anthesis. *Physiol. Mol. Plant* 63, 271–279. doi: 10.1016/j.pmpp.2004.02.001
- Little, C. R., Perumal, R., Tesso, T., Prom, L. K., Odvody, G. N., and Magill, C. W. (2011). Sorghum Pathology and Biotechnology -A Fungal Disease Perspective: Part I. Grain Mold, Head Smut, and Ergot. *The European Journal of Plant Science and Biotechnology* *vol page,
- Liu, H., Du, Y., Chu, H., Shih, C. H., Wong, Y. W., Wang, M., et al. (2010). Molecular Dissection of the Pathogen-Inducible 3-Deoxyanthocyanidin Biosynthesis Pathway in Sorghum. *Plant Cell Physiol* 51, 1173–1185. doi: 10.1093/pcp/pcq080
- Logrieco, A., Moretti, A., and Solfrizzo, M. (2009). Alternaria toxins and plant diseases: an overview of origin, occurrence and risks. *World Mycotoxin J* 2, 129–140. doi: 10.3920/wmj2009.1145
- Lombard, M. J. (2014). Mycotoxin Exposure and Infant and Young Child Growth in Africa: What Do We Know? *Ann Nutr Metab* 64, 42–52. doi: 10.1159/000365126
- Mahato, D. K., Lee, K. E., Kamle, M., Devi, S., Dewangan, K. N., Kumar, P., et al. (2019). Aflatoxins in Food and Feed: An Overview on Prevalence, Detection and Control Strategies. *Front Microbiol* 10:2266. doi: 10.3389/fmicb.2019.02266
- Mareya, C. R., Tugizimana, F., Piater, L. A., Madala, N. E., Steenkamp, P. A., and Dubery, I. A. (2019). Untargeted Metabolomics Reveal Defense-Related Metabolic Reprogramming in Sorghum bicolor against Infection by *Burkholderia andropogonis*. *Metabolites* 9, 8–8. doi: 10.3390/metabo9010008
- Marin, S., Ramos, A. J., Cano-Sancho, G., and Sanchis, V. (2013). Mycotoxins: Occurrence, toxicology, and exposure assessment. *Food Chem Toxicol* 60, 218–237. doi: 10.1016/j.fct.2013.07.047
- Martens, S., and Mithöfer, A. (2005). Flavones and flavone synthases. *Phytochemistry* 66, 2399–2407. doi: 10.1016/j.phytochem.2005.07.013
- Melake-Berhan, A., Butler, L. G., Ejeta, G., and Menkir, A. (1996). Grain Mold Resistance and Polyphenol Accumulation in Sorghum. *J Agr Food Chem* 44, 2428–2434. doi: 10.1021/jf950580x
- Menkir, A., Ejeta, I. G., Butler, L., and Melakeberhan, A. (1996). Physical and Chemical Kernel Properties Associated with Resistance to Grain Mold in Sorghum. *Cereal Chemistry* 73, 617.

- Meseka, S., Williams, W. P., Warburton, M. L., Brown, R. L., Augusto, J., Ortega-Beltran, A., et al. (2018). Heterotic affinity and combining ability of exotic maize inbred lines for resistance to aflatoxin accumulation. *Euphytica* 214, 184. doi: 10.1007/s10681-018-2254-8
- Mizuno, H., Yazawa, T., Kasuga, S., Sawada, Y., Kanamori, H., Ogo, Y., et al. (2016). Expression of Flavone Synthase II and Flavonoid 3'-Hydroxylase Is Associated with Color Variation in Tan-Colored Injured Leaves of Sorghum. *Front Plant Sci* 7:1718. doi: 10.3389/fpls.2016.01718
- Mizuno, H., Yazawa, T., Kasuga, S., Sawada, Y., Ogata, J., Ando, T., et al. (2014). Expression level of a flavonoid 3'-hydroxylase gene determines pathogen-induced color variation in sorghum. *Bmc Res Notes* 7:761. doi: 10.1186/1756-0500-7-761
- Moretti, A., Bennett, G. A., Logrieco, A., Bottalico, A., and Beremand, M. N. (1995). Fertility of *Fusarium moniliforme* from maize and sorghum related to fumonisin production in Italy. *Mycopathologia* 131, 25–29. doi: 10.1007/bf01103900
- Mousa, W., Ghazali, F. M., Jinap, S., Ghazali, H. M., and Radu, S. (2011). Modelling the effect of water activity and temperature on growth rate and aflatoxin production by two isolates of *Aspergillus flavus* on paddy. *J Appl Microbiol* 111, 1262–1274. doi: 10.1111/j.1365-2672.2011.05134.x
- Mpofu, L. T., and McLaren, N. W. (2014). Ergosterol concentration and variability in genotype-by-pathogen interaction for grain mold resistance in sorghum. *Planta* 240, 239–250. doi: 10.1007/s00425-014-2081-7
- Mukherjee, K., and Lakshminarasimham, A. V. (1995). Aflatoxin contamination of sorghum seeds during storage under controlled conditions. *Zentralblatt Für Bakteriologie* 282, 237–243. doi: 10.1016/s0934-8840(11)80123-4
- Munkvold, G. P. (2003). Epidemiology of *Fusarium* Diseases and their Mycotoxins in Maize Ears. *Eur J Plant Pathol* 109, 705–713. doi: 10.1023/a:1026078324268
- Nelson, P. E., Plattner, R. D., Shackelford, D. D., and Desjardins, A. E. (1991). Production of fumonisins by *Fusarium moniliforme* strains from various substrates and geographic areas. *Appl Environ Microb* 57, 2410–2412.
- Nelson, P. E., Plattner, R. D., Shackelford, D. D., and Desjardins, A. E. (1992). Fumonisin B1 production by *Fusarium* species other than *F. moniliforme* in section *Liseola* and by some related species. *Appl Environ Microb* 58, 984–989.
- Nicholson, R. L., Kollipara, S. S., Vincent, J. R., Lyons, P. C., and Cadena-Gomez, G. (1987). Phytoalexin synthesis by the sorghum mesocotyl in response to infection by pathogenic and nonpathogenic fungi. *Proc National Acad Sci* 84, 5520–5524. doi: 10.1073/pnas.84.16.5520
- Nida, H., Girma, G., Mekonen, M., Lee, S., Seyoum, A., Desalegn, K., et al. (2019). Identification of sorghum grain mold resistance loci through genome wide association mapping. *J Cereal Sci* 85, 295–304. doi: 10.1016/j.jcs.2018.12.016
- Nida, H., Girma, G., Mekonen, M., Tirfessa, A., Seyoum, A., Bejiga, T., et al. (2021). Genome-wide association analysis reveals seed protein loci as determinants of variations in grain mold resistance in sorghum. *Theor Appl Genet* 134, 1–18. doi: 10.1007/s00122-020-03762-2
- Nithya, R., Sharma, R., Rao, V. P., Gopalakrishnan, S., and Thakur, R. P. (2013). Biochemical characterisation of grain mould resistant and susceptible genotypes and PGPR-induced resistance in the host to *Curvularia lunata* and *Fusarium proliferatum*. *Archives Phytopathology Plant Prot* 46, 980–989. doi: 10.1080/03235408.2012.755824
- Oliveira, R. C., Goncalves, S. S., Silva, C. D. C., Dilkin, P., Madrid, H., and Correa, B. (2019). Polyphasic characterization of *Epicoccum sorghinum*: a tenuazonic acid producer isolated from sorghum grain. *Int. J. Food Microbiol.* 292, 1–7. doi: 10.1016/j.ijfoodmicro.2018.12.004
- Omotayo, O. P., Omotayo, A. O., Mwanza, M., and Babalola, O. O. (2019). Prevalence of Mycotoxins and Their Consequences on Human Health. *Toxicol Res* 35, 1–7. doi: 10.5487/tr.2019.35.1.001
- Ostry, V., Malir, F., Toman, J., and Grosse, Y. (2017). Mycotoxins as human carcinogens—the IARC Monographs classification. *Mycotoxin Res* 33, 65–73. doi: 10.1007/s12550-016-0265-7
- Park, D. L., Njapau, H., and Boutrif, E. (1999). *Minimizing risks posed by mycotoxins utilizing the HACCP concept*. Rome: FAO Food, 23.
- Pasha, I., Riaz, A., Saeed, M., and Randhawa, M. A. (2015). Exploring the Antioxidant Perspective of Sorghum and Millet. *J Food Process Pres* 39, 1089–1097. doi: 10.1111/jfpp.12323
- Patriarca, A., Azcarate, M. P., Terminiello, L., and Pinto, V. F. (2007). Mycotoxin production by *Alternaria* strains isolated from Argentinean wheat. *Int. J. Food Microbiol.* 119, 219–222. doi: 10.1016/j.ijfoodmicro.2007.07.055
- Pena, G. A., Cavaglieri, L. R., and Chulze, S. N. (2019). *Fusarium* species and moniliformin occurrence in sorghum grains used as ingredient for animal feed in Argentina. *J Sci Food Agr* 99, 47–54. doi: 10.1002/jsfa.9140
- Pfiegler, W. P., Pócsi, I., Györi, Z., and Pusztahelyi, T. (2020). The *Aspergilli* and Their Mycotoxins: Metabolic Interactions With Plants and the Soil Biota. *Front Microbiol* 10:2921. doi: 10.3389/fmicb.2019.02921
- Poloni, A., Schirawski, J., Poloni, A., and Schirawski, J. (2014). Red Card for Pathogens: Phytoalexins in Sorghum and Maize. *Molecules* 19, 9114–9133. doi: 10.3390/molecules19079114
- Prom, L. K. (2004). The Effects of *Fusarium thapsinum*, *Curvularia lunata*, and Their Combination on Sorghum Germination and Seed Mycoflora. *J New Seeds* 6, 39–49. doi: 10.1300/j153v06n01_03
- Prom, L. K., Cisse, N., Perumal, R., and Cuevas, H. (2017). Screening of Sorghum Lines Against Long Smut and Grain Mold Pathogens. *International Journal of Plant Pathology* 8, 23–27.
- Prom, L. K., Cuevas, H., Isakeit, T., and Magill, C. (2020b). Screening Sorghum Accessions for Resistance against Anthracnose and Grain Mold through Inoculating with Pathogens. *J Exp Agric Int* 42, 73–83. doi: 10.9734/jeai/2020/v42i130453
- Prom, L. K., Cuevas, H. E., Ahn, E., Isakeit, T., Rooney, W. L., and Magill, C. (2020a). Genome-wide association study of grain mold resistance in sorghum association panel as affected by inoculation with *Alternaria alternata* alone and *Alternaria alternata*, *Fusarium thapsinum*, and *Curvularia lunata* combined. *Eur J Plant Pathol* 157, 783–798. doi: 10.1007/s10658-020-02036-3
- Prom, L. K., Isakeit, T., Perumal, R., Erpelding, J. E., Rooney, W., and Magill, C. W. (2011). Evaluation of the Ugandan sorghum accessions for grain mold and anthracnose resistance. *Crop Prot.* 30, 566–571. doi: 10.1016/j.cropro.2010.12.025
- Prom, L. K., Perumal, R., Cissé, N., and Little, C. R. (2014). Evaluation of Selected Sorghum Lines and Hybrids for Resistance to Grain Mold and Long Smut Fungi in Senegal. *West Africa. Plant Heal Prog* 15, 74–77. doi: 10.1094/php-rs-13-0128
- Prom, L. K., Radwan, G., Perumal, R., Cuevas, H., Katile, S. O., Isakeit, T., et al. (2016). Grain Biodeterioration of Sorghum Converted Lines Inoculated with a Mixture of *Fusarium thapsinum* and *Curvularia lunata*. *Plant Pathology J* 16, 19–24. doi: 10.3923/ppj.2017.19.24
- Prom, L. K., Waniska, R. D., Kollo, A. I., and Rooney, W. L. (2003). Response of eight sorghum cultivars inoculated with *Fusarium thapsinum*, *Curvularia lunata*, and a mixture of the two fungi. *Crop Prot* 22, 623–628. doi: 10.1016/s0261-2194(02)00246-6
- Rajini, S. B., Nandhini, M., Udayashankar, A. C., Niranjana, S. R., Lund, O. S., and Prakash, H. S. (2020). Diversity, plant growth—promoting traits, and biocontrol potential of fungal endophytes of Sorghum bicolor. *Plant Pathol* 69, 642–654. doi: 10.1111/ppa.13151
- Ramakrishna, Y., Bhat, R. V., and Ravindranath, V. (1989). Production of deoxynivalenol by *Fusarium* isolates from samples of wheat associated with a human mycotoxicosis outbreak and from sorghum cultivars. *Appl Environ Microb* 55, 2619–2620.
- Ratnavathi, C., and Sashidhar, R. (2004). Induction of chitinase in response to *Aspergillus* infection in sorghum (*Sorghum bicolor* (L.) Moench). *J Sci Food Agr* 84, 1521–1527. doi: 10.1002/jsfa.1799
- Razzaq, A., Sadia, B., Raza, A., Hameed, M. K., and Saleem, F. (2019). Metabolomics: A Way Forward for Crop Improvement. *Metabolites* 9, 303. doi: 10.3390/metabo9120303
- Reddy, B. V. S. (2000). Breeding Mold Resistant Sorghum Cultivars. *Technical and Institutional Options for Sorghum Grain Mold Management* 195–224. *vol,
- Reddy, K. R. N., Raghavender, C. R., Salleh, B., Reddy, C. S., and Reddy, B. N. (2011). Potential of aflatoxin B1 production by *Aspergillus flavus* strains on commercially important food grains. *Int J Food Sci Technology* 46, 161–165. doi: 10.1111/j.1365-2621.2010.02468.x
- Righini, S., Rodriguez, E. J., Berosich, C., Grotewold, E., Casati, P., and Ferreyra, M. L. F. (2019). Apigenin produced by maize flavone synthase I and II protects plants against UV-B—induced damage. *Plant Cell Environ* 42, 495–508. doi: 10.1111/pce.13428
- Rodriguez-Herrera, R., Rooney, W. L., Frederiksen, D. T. R., and Frederiksen, R. A. (2000). Inheritance of Grain Mold Resistance in Grain Sorghum without a Pigmented Testa. *Environ Health Persp* 109, 239–243. doi: 10.1289/ehp.01109s2239

- Rooney, L. W., and Murty, D. S. (1982). "Evaluation of Sorghum Food Quality. Sorghum in the Eighties," in *Proceedings of the International Symposium on Sorghum*, (Telangana).
- Sarma, U. P., Bhetaria, P. J., Devi, P., and Varma, A. (2017). Aflatoxins: Implications on Health. *Indian J Clin Biochem* 32, 124–133. doi: 10.1007/s12291-017-0649-2
- Schmale, D. G. III, and Munkvold, G. P. (2009). Mycotoxins in Crops: A Threat to Human and Domestic Animal Health. *Plant Heal Instr.* doi: 10.1094/phi-i-2009-0715-01 *vol page,
- Schöneberg, T., Kibler, K., Sulyok, M., Musa, T., Bucheli, T. D., Mascher, F., et al. (2018). Can plant phenolic compounds reduce Fusarium growth and mycotoxin production in cereals? *Food Addit Contam Part 35*, 1–16. doi: 10.1080/19440049.2018.1538570
- Schutt, C., and Netzky, D. (1991). Effect of apiforol and apigeninidin on growth of selected fungi. *J Chem Ecol* 17, 2261–2266. doi: 10.1007/bf00988006
- Seetharaman, K., Whitehead, E., Keller, N. P., Waniska, R. D., and Rooney, L. W. (1997). In Vitro Activity of Sorghum Seed Antifungal Proteins against Grain Mold Pathogens. *J Agr Food Chem* 45, 3666–3671. doi: 10.1021/jf960924z
- Seifert, K. A., Aoki, T., Baayen, R. P., Brayford, D., Burgess, L. W., Chulze, S., et al. (2003). The Name Fusarium Moniliforme Should no Longer be Used. *Mycol Res* 107, 643–644. doi: 10.1017/s095375620323820x
- Sharma, R., Upadhyaya, H. D., Thakur, V. G. R., Thakur, R. P., and Rao, V. P. (2010). Resistance to Grain Mold and Downy Mildew in a Mini-Core Collection of Sorghum Germplasm. *Plant Dis.* 94, 439–444.
- Shen, S., Huang, R., Li, C., Wu, W., Chen, H., Shi, J., et al. (2018). Phenolic Compositions and Antioxidant Activities Differ Significantly among Sorghum Grains with Different Applications. *Molecules* 23, 1203. doi: 10.3390/molecules23051203
- Snyder, B. A., and Nicholson, R. L. (1990). Synthesis of Phytoalexins in Sorghum as a Site-Specific Response to Fungal Ingress. *Science* 248, 1637–1639. doi: 10.1126/science.248.4963.1637
- Somashekar, D., Rati, E. R., Anand, S., and Chandrashekar, A. (2004). Isolation, enumeration and PCR characterization of aflatoxigenic fungi from food and feed samples in India. *Food Microbiology* 21, 809–813.
- Soni, P., Gangurde, S. S., Ortega-Beltran, A., Kumar, R., Parmar, S., Sudini, H. K., et al. (2020). Functional Biology and Molecular Mechanisms of Host-Pathogen Interactions for Aflatoxin Contamination in Groundnut (*Arachis hypogaea* L.) and Maize (*Zea mays* L.). *Front Microbiol* 11:227. doi: 10.3389/fmicb.2020.00227
- Stuper-Szablewska, K., and Perkowski, J. (2017). Phenolic acids in cereal grain: Occurrence, biosynthesis, metabolism and role in living organisms. *Crit Rev Food Sci* 59, 00–00. doi: 10.1080/10408398.2017.1387096
- Taleon, V., Dykes, L., Rooney, W. L., and Rooney, L. W. (2012). Effect of genotype and environment on flavonoid concentration and profile of black sorghum grains. *J Cereal Sci* 56, 470–475. doi: 10.1016/j.jcs.2012.05.001
- Thakur, R., Rao, V., and Sharma, R. (2008). Characterization of grain mold resistant sorghum germplasm accessions for physio-morphological traits. *Open Access Journal published by ICRISAT* 6, 1–7.
- Tohge, T., Watanabe, M., Hoefgen, R., and Fernie, A. R. (2013). Shikimate and Phenylalanine Biosynthesis in the Green Lineage. *Front Plant Sci* 4:62. doi: 10.3389/fpls.2013.00062
- Tola, M., and Kebede, B. (2016). Occurrence, importance and control of mycotoxins: A review. *Cogent Food Agric* 2, 1191103. doi: 10.1080/23311932.2016.1191103
- Torre-Hernandez, M. E., de la, Vicente, M. R.-S., Greaves-Fernandez, N., Cruz-Ortega, R., and Plasencia, J. (2010). Fumonisin B1 induces nuclease activation and salicylic acid accumulation through long-chain sphingoid base build-up in germinating maize. *Physiol Mol Plant* 74, 337–345. doi: 10.1016/j.pmp.2010.05.004
- Tralamazza, S. M., Piacentini, K. C., Iwase, C. H. T., and Rocha, L. O. (2018). Toxigenic alternaria species: impact in cereals worldwide. *Curr. Opin. Food Sci.* 23, 57–63. doi: 10.1016/j.cofs.2018.05.002
- Turner, M. F., Heuberger, A. L., Kirkwood, J. S., Collins, C. C., Wolfrum, E. J., Broeckling, C. D., et al. (2016). Non-targeted Metabolomics in Diverse Sorghum Breeding Lines Indicates Primary and Secondary Metabolite Profiles Are Associated with Plant Biomass Accumulation and Photosynthesis. *Front Plant Sci* 7:953. doi: 10.3389/fpls.2016.00953
- Upadhyaya, H. D., Pundir, R. P. S., Dwivedi, S. L., Gowda, C. L. L., Reddy, V. G., and Singh, S. (2009). Developing a Mini Core Collection of Sorghum for Diversified Utilization of Germplasm. *Crop Sci* 49, 1769–1780. doi: 10.2135/cropsci2009.01.0014
- USDA (2016). *Grain Fungal Diseases and Mycotoxin Reference*.
- USDA-ARS (2018b). *What are Mycotoxins?* Available online at: <https://www.ars.usda.gov/midwest-area/peoria-il/national-center-for-agricultural-utilization-research/mycotoxin-prevention-and-applied-microbiology-research/docs/what-are-mycotoxins> (accessed October 18, 2020).
- USDA-ARS (2018a). *Mycotoxin Prevention and Applied Microbiology Research*. Peoria, IL. Available online at: <https://www.ars.usda.gov/midwest-area/peoria-il/national-center-for-agricultural-utilization-research/mycotoxin-prevention-and-applied-microbiology-research/docs/what-are-mycotoxins> (accessed October 18, 2020).
- Usha, C. M., Patkar, K. L., Shetty, H. S., Kennedy, R., and Lacey, J. (1994). Mycoflora of developing sorghum grains with special reference to *Aspergillus flavus*. *Tropical Science* 34, 360.
- Verwaaijen, B., Wibberg, D., Winkler, A., Zrenner, R., Bednarz, H., Niehaus, K., et al. (2019). A comprehensive analysis of the *Lactuca sativa*, L. transcriptome during different stages of the compatible interaction with *Rhizoctonia solani*. *Sci Rep-uk* 9, 7221. doi: 10.1038/s41598-019-43706-5
- Visconti, A., and Doko, M. B. (1994). Survey of fumonisin production by *Fusarium* isolated from cereals in Europe. *J. Aoac. Int.* 77, 546–550. doi: 10.1093/jaoac/77.2.546
- Vismer, H. F., Shephard, G. S., van der Westhuizen, L., Mngqawa, P., Bushula-Njah, V., and Leslie, J. F. (2019). Mycotoxins produced by *Fusarium proliferatum* and *F. pseudonygamai* on maize, sorghum and pearl millet grains in vitro. *Int. J. Food Microbiol.* 296, 31–36. doi: 10.1016/j.jfoodmicro.2019.02.016
- Waniska, R. (2000). "Structure, Phenolic Compounds, and Antifungal Proteins of Sorghum Caryopses," in *Technical and Institutional Options for Sorghum Grain Mold Management: Proceedings of an International Consultation*, eds A. Chandrashekar, R. Bandyopadhyay, and A. J. Hall (Patancheru: ICRISAT), 18–19.
- Waniska, R. D., Venkatesha, R. T., Chandrashekar, A., Krishnaveni, S., Bejosano, F. P., Jeoung, J., et al. (2001). Antifungal Proteins and Other Mechanisms in the Control of Sorghum Stalk Rot and Grain Mold. *J Agr Food Chem* 49, 4732–4742. doi: 10.1021/jf010007f
- Ward, T. J., Clear, R. M., Rooney, A. P., O'Donnell, K., Gaba, D., Patrick, S., et al. (2008). An adaptive evolutionary shift in *Fusarium* head blight pathogen populations is driving the rapid spread of more toxigenic *Fusarium graminearum* in North America. *Fungal Genet Biol* 45, 473–484. doi: 10.1016/j.fgb.2007.10.003
- Wild, C. P., and Gong, Y. Y. (2010). Mycotoxins and human disease: a largely ignored global health issue. *Carcinogenesis* 31, 71–82. doi: 10.1093/carcin/bgp264
- Willett, K. M., Dunn, R. J. H., Thorne, P. W., Bell, S., Podesta, M., de Parker, D. E., et al. (2014). HadISDH land surface multi-variable humidity and temperature record for climate monitoring. *Clim Past* 10, 1983–2006. doi: 10.5194/cp-10-1983-2014
- Williams, R. J., and Rao, K. N. (1981). A Review of Sorghum Grain Moulds. *Trop Pest Manage* 27, 200–211. doi: 10.1080/09670878109413652
- Winn, R. T., and Lane, G. T. (1978). Aflatoxin Production on High Moisture Corn and Sorghum with a Limited Incubation1. *J Dairy Sci* 61, 762–764. doi: 10.3168/jds.s0022-0302(78)83645-5
- Wu, Y., Li, X., Xiang, W., Zhu, C., Lin, Z., Wu, Y., et al. (2012). Presence of tannins in sorghum grains is conditioned by different natural alleles of Tannin1. *Proc National Acad Sci* 109, 10281–10286. doi: 10.1073/pnas.1201700109
- Xiong, Y., Zhang, P., Warner, R. D., and Fang, Z. (2019a). 3-Deoxyanthocyanidin Colorant: Nature, Health, Synthesis, and Food Applications. *Compr Rev Food Sci F* 18, 1533–1549. doi: 10.1111/1541-4337.12476
- Xiong, Y., Zhang, P., Warner, R. D., and Fang, Z. (2019b). Sorghum Grain: From Genotype, Nutrition, and Phenolic Profile to Its Health Benefits and Food Applications. *Compr Rev Food Sci F* 18, 2025–2046. doi: 10.1111/1541-4337.12506
- Yonekura-Sakakibara, K., Higashi, Y., and Nakabayashi, R. (2019). The Origin and Evolution of Plant Flavonoid Metabolism. *Front Plant Sci* 10:943. doi: 10.3389/fpls.2019.00943

- Zhang, Y., Ngu, D. W., Carvalho, D., Liang, Z., Qiu, Y., Roston, R. L., et al. (2017a). Differentially Regulated Orthologs in Sorghum and the Subgenomes of Maize. *Plant Cell* 29, 1938–1951. doi: 10.1105/tpc.17.00354
- Zhang, Y., Zhao, L., Zhao, J., Li, Y., Wang, J., Guo, R., et al. (2017b). S5H/DMR6 Encodes a Salicylic Acid 5-Hydroxylase That Fine-Tunes Salicylic Acid Homeostasis. *Plant Physiol* 175, 1082–1093. doi: 10.1104/pp.17.00695
- Zhao, R., Bean, S. R., Wang, D., Park, S. H., Schober, T. J., and Wilson, J. D. (2009). Small-scale mashing procedure for predicting ethanol yield of sorghum grain. *J Cereal Sci* 49, 230–238. doi: 10.1016/j.jcs.2008.10.006

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ackerman, Wenndt and Boyles. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Identifying Anti-Oxidant Biosynthesis Genes in Pearl Millet [*Pennisetum glaucum* (L.) R. Br.] Using Genome—Wide Association Analysis

Chandra Bhan Yadav^{1†}, Jayanti Tokas^{1,2†}, Devvart Yadav^{1,2†}, Ana Winters¹, Ram B. Singh³, Rama Yadav¹, Prakash I. Gangashetty⁴, Rakesh K. Srivastava³ and Rattan S. Yadav^{1*}

OPEN ACCESS

Edited by:

Wolfram Weckwerth,
University of Vienna, Austria

Reviewed by:

Arindam Ghatak,
University of Vienna, Austria
Paola Leonetti,
National Research Council (CNR), Italy

*Correspondence:

Rattan S. Yadav
rsy@aber.ac.uk

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 27 August 2020

Accepted: 19 April 2021

Published: 28 May 2021

Citation:

Yadav CB, Tokas J, Yadav D,
Winters A, Singh RB, Yadav R,
Gangashetty PI, Srivastava RK and
Yadav RS (2021) Identifying
Anti-Oxidant Biosynthesis Genes in
Pearl Millet [*Pennisetum glaucum* (L.)
R. Br.] Using Genome—Wide
Association Analysis.
Front. Plant Sci. 12:599649.
doi: 10.3389/fpls.2021.599649

¹ Institute of Biological Environmental and Rural Sciences (IBERS), Aberystwyth University, Aberystwyth, United Kingdom,

² Chaudhary Charan Singh (CCS) Haryana Agricultural University, Hisar, India, ³ International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India, ⁴ International Crops Research Institute for the Semi-Arid Tropics, Niamey, Niger

Pearl millet [*Pennisetum glaucum* (L.) R Br.] is an important staple food crop in the semi-arid tropics of Asia and Africa. It is a cereal grain that has the prospect to be used as a substitute for wheat flour for celiac patients. It is an important antioxidant food resource present with a wide range of phenolic compounds that are good sources of natural antioxidants. The present study aimed to identify the total antioxidant content of pearl millet flour and apply it to evaluate the antioxidant activity of its 222 genotypes drawn randomly from the pearl millet inbred germplasm association panel (PMiGAP), a world diversity panel of this crop. The total phenolic content (TPC) significantly correlated with DPPH (1,1-diphenyl-2-picrylhydrazyl) radical scavenging activity (% inhibition), which ranged from 2.32 to 112.45% and ferric-reducing antioxidant power (FRAP) activity ranging from 21.68 to 179.66 (mg ascorbic acid eq./100g). Genome-wide association studies (GWAS) were conducted using 222 diverse accessions and 67 K SNPs distributed across all the seven pearl millet chromosomes. Approximately, 218 SNPs were found to be strongly associated with DPPH and FRAP activity at high confidence [$-\log(p) > 3.0-7.4$]. Furthermore, flanking regions of significantly associated SNPs were explored for candidate gene harvesting. This identified 18 candidate genes related to antioxidant pathway genes (flavanone 7-O-beta-glycosyltransferase, GDGL esterase/lipase, glutathione S-transferase) residing within or near the association signal that can be selected for further functional characterization. Patterns of genetic variability and the associated genes reported in this study are useful findings, which would need further validation before their utilization in molecular breeding for high antioxidant-containing pearl millet cultivars.

Keywords: pearl millet [*Pennisetum glaucum* (L.) R. Br.], antioxidant activity, phenolics, germplasm, genome-wide association study, marker-trait associations, candidate genes

INTRODUCTION

Pearl millet [*Pennisetum glaucum* (L.) R Br.] provides nutritious staple food grains worldwide especially to the poorest rural households in some of the hottest and driest rainfed farming regions of Africa and Asia. Global cereal production has been estimated to increase by ~1 billion tons over the last 50 years (FAOSTAT database, 2017). Pearl millet grown on 33M ha worldwide is one of the four most important cereals grown in the tropical semi-arid regions of the world. It has excellent sustainability credentials, as the crop can easily survive on marginal lands, in harsh climatic conditions, and has a short growing season to complete its life cycle. It also is well-suited to multiple cropping systems under irrigated and dry land farming. It contains on average 92.5% dry matter, 2.1% ash, 2.8% crude fiber, 7.8% crude fat, 13.6% crude protein, and 63.2% starch, mostly of the resistant form, and high iron levels. It has a high energy content, lower readily available starch levels, higher fiber (1.2 g/100g, most of which is insoluble), and 8–15 times greater α -amylase activity compared with wheat. Pearl millet also has a low glycemic index (50) and is gluten-free, thus, an ideal candidate grain for use in the functional-food market worldwide (Ali et al., 2003; Ragaei et al., 2006; Saleh et al., 2013). It also contains several phenolic compounds such as benzoic and cinnamic acid derivatives, anthocyanidins, flavonoids, lignans, and phytoestrogens, which play an important role in disease prevention (Muthamilarasan et al., 2016).

There are growing scientific evidences available, which confirm that diets, especially the plant food-based diets, rich in antioxidant compounds, offer a lower risk of developing cardiovascular diseases, certain kinds of cancers, and age-related degenerative processes (Pandey and Rizvi, 2009; Shahidi and Chandrasekara, 2013). Particular attention has been drawn to the role they play as “free radical scavengers,” which has provoked numerous studies into studying phenolic compounds in many plants, including cereals. The most studied sources of natural antioxidants are vegetables, fruits, and cereals (Khan et al., 2015). The presence of antioxidants in cereals is a consequence of the fact that all biological systems, including cereals, have a natural tendency to minimize the destructive potential of oxidation reactions, and consequently, they have developed their own multifunctional defense systems (Shahidi, 2000). Antioxidants found in cereals have the advantage of maintaining antioxidant capacity inside the human body following consumption (Serea and Barna, 2011). Cereals, especially millets, are the most commonly consumed foods in India and sub-Saharan Africa. They contain a wide range of phenolic compounds that are good sources of natural antioxidants (Ilesanmi and Akinloye, 2013). Antioxidant capacity of different cereal products, such as corn, wheat, rice, oats, and ready-to-eat breakfast cereals has been reported previously (Adom and Liu, 2002; Choi et al., 2007; Singh et al., 2012). In these cereals, antioxidant activity of different extracts correlated very well with their total plant phenolic content (Zieliński and Kozłowska, 2000). However, information on the extent of genetic variability and genes involved for antioxidant activities in pearl millet germplasm is scanty.

Rapid human population growth on a global scale is boosting the demand for a corresponding increase in crop grain yield, coupled with better nutritional credentials for a food secure future. Understanding the molecular and genetic control of useful traits, such as yield and nutritional quality, therefore, is becoming a major objective in the genetic study of staple cereal crops (Jin et al., 2010). Association mapping has been demonstrated as a powerful tool for dissecting genetic control of agronomic and other traits in plants (Yu and Buckler, 2006; Buckler and Gore, 2007; Zhu et al., 2008; Sehgal et al., 2015). Association panels today exist for a number of key crops (for example maize, wheat, sorghum, and pearl millet) that are routinely used in performing association analysis (Flint-Garcia et al., 2005; Maccaferri et al., 2006; Casa et al., 2008; Varshney et al., 2017). In pearl millet, an association panel known as pearl millet inbred germplasm association panel (PMiGAP) has been assembled from the world collection of pearl millet germplasm maintained at ICRISAT (Sehgal et al., 2015), which has recently been re-sequenced making available 28 million SNPs (Varshney et al., 2017) for use in genome-wide association studies. Availability of such resources is paving a powerful way for linkage disequilibrium (LD) mapping and association analysis of traits in pearl millet.

In the present study, we report on genetic variation for antioxidant-related traits in pearl millet, and the possible candidate genes associated with such traits, using genome-wide association studies.

MATERIALS AND METHODS

Pearl Millet Germplasm Association Panel's Genotypes Used

A random set of 222 genotypes from within the pearl millet germplasm association panel (PMiGAP) was included in this study for antioxidant analysis (**Supplementary Table 1**). The PMiGAP is a collection of genotypes drawn from within the pearl millet core collection, landraces, cultivars, and breeding lines as explained in Sehgal et al. (2015) and represents geographical regions from 23 pearl millet growing countries across the world. The PMiGAP genotypes, thus, included in this study represent the entire global diversity of cultivated pearl millet. Further details of the origin of each of the PMiGAP genotypes used in the study are available in Sehgal et al. (2015) and Varshney et al. (2017). Seeds of each of the PMiGAP genotypes used for antioxidant determination were bulk multiplied by growing them in a uniform field condition following standard agronomic and seed multiplication procedure (Upadhyaya et al., 2008; Ramya et al., 2018). Each entry was planted in three rows by maintaining 15 cm between plants and 75 cm between rows. Field was applied with 100 kg/ha of DAP (di-ammonium phosphate) as basal dose; thinning was done to one plant per hill. Weeding was done two times during the seed multiplication plot. Each individual head was selfed before the emergence of panicle, and strict pollination was controlled to get the pure seeds of each line. Harvesting was done during physiological maturity. Threshing was done with

Winterstieger single head thresher and cleaned to remove the remaining debris.

Preparation of Extracts for Antioxidant Analysis From Pearl Millet Flour

Pearl millet flour was extracted for antioxidant analysis according to a modification of the method described in Akpanika et al. (2017). A quantity of washed grains (1 g) was milled in a customized robotic instrument (Labmann Automation, Middlesbrough, UK) to obtain a fine flour. Thirty milligrams of flour was extracted with 5 ml of 70% ethanol with intermittent stirring for 1 h. The extract was then centrifuged at $10,000 \times g$ for 10 min at room temperature, and the supernatant was transferred to clean tubes. The pellets were then washed with 2.0 ml of 70% ethanol and centrifuged again at $10,000 \times g$ for 10 min. The second supernatant fraction was added to the first to maximize recovery of target compounds. The ethanol was then removed using a heated centrifugal evaporator (Jouan RC10.22, Saint Herblain, France) set to 70°C . The dried extract was re-dissolved with 70% ethanol to an equivalent concentration of 15 mg/ml of the original sample and stored at 4°C until analysis. All biochemical analyses (see below) were carried out in triplicate, which allowed us to calculate the repeatability of these measurements.

Determination of 1,1-Diphenyl-2-Picrylhydrazyl Radical Scavenging Activity

DPPH is a stable free radical for the determination of antioxidant or radical scavenging capabilities. The DPPH radical portion of the molecule is a centralized nitrogen atom that gives rise to a maximum absorbance at 515 nm in methanol in its oxidized form. When a solution of DPPH in its radical form is mixed with a proton-donating substance such as antioxidants, the radicals are scavenged, and $\text{DPPH}=\text{H}$ is formed with a concomitant decrease in absorbance. The ability of the pearl millet ethanolic extracts to scavenge free radicals was determined against a very stable free radical DPPH (1,1-diphenyl-2-picrylhydrazyl). DPPH is violet in color, while the reduced product is colorless, and the loss of color was determined in a plate reader (Hauck, 2018). Aliquots of the sample extract (50 μl) at different concentrations were added to ethanolic solutions of 500 μl of DPPH (0.12 mg/ml). Each mixture was left for 20 min at room temperature in the dark. The absorbance was measured at 517 nm, and the activity was expressed as percentage of DPPH radical relative to control using the following equation: $\text{DPPH scavenging activity (\%)} = [(\text{absorbance of control} - \text{absorbance of sample}) / \text{absorbance of control}] \times 100$.

Ferric-Reducing Antioxidant Power Assay

FRAP assay is a novel method for assessing antioxidant power where the ferric-reducing ability of a sample extract is tested. Ferric to ferrous ion reduction at low pH causes a colored ferrous-tripyridyltriazine complex to form. FRAP values are obtained by comparing the absorbance change at 593 nm in test reaction mixtures with those containing ferrous ions in known

concentration (Benzie and Strain, 1999). The ferric-reducing antioxidant power (FRAP) assay measures the antioxidant potential of antioxidants to reduce the Fe^{3+} /tripyridyl-s-triazine complex present in stoichiometric excess to the blue ferrous form (Hauck, 2018). FRAP reagent was freshly prepared by mixing together 10 mM 2,4,6-tripyridyl-s-triazine (TPTZ) and 20 mM ferric chloride in 0.25 M acetate buffer, pH 3.6 in proportions of 1:1:10 (v/v), respectively. Pearl millet extract (50 μl) was added to 1.5 ml of FRAP reagent. The absorbance was read at 593 nm after 4-min incubation at ambient temperature against distilled water as a blank in a plate reader. A calibration curve of ascorbic acid concentration (100–1,000 $\mu\text{mol/L}$) versus absorbance was used to calculate values, and results are expressed in mg/ml ascorbic equivalents/g dry weight for plant total extracts from three determinations.

Sample Separation by High-Performance Liquid Chromatography and Statistical Analysis

Among the analyzed pearl millet samples, 40 samples were picked from the higher, lower, and mid-range antioxidant potential, and evaluated for total phenolic and flavonoid content by high-performance liquid chromatography with online photodiode array detection and electrospray ionization-ion trap tandem mass spectrometry (HPLC-PDA-ESI/MSⁿ) analysis (Supplementary Table 2). Samples were prepared for analysis by solid phase extraction with Waters Sep-Pak C18 500 mg cartridges (WAT036945) prepared by passing through 4 ml of 100% methanol followed by 4 ml of purified water. The sample (500 μl) was then loaded onto the cartridge. The unbound aqueous fraction was eluted with 4 ml of purified water, and the bound fraction was then eluted with 4 ml methanol. Both fractions were collected in separate clean vials. The samples were then fully dried using a heated centrifugal evaporator at 13,000 rpm and 70°C , and reconstituted in 100 μl of 70% methanol. Secondary metabolites were analyzed by reverse-phase HPLC-PDA-ESI/MSⁿ.

The analyses were carried out on a Thermo Finnigan LC-MS system (Thermo Electron Corp, Waltham, MA, USA) with a Finnigan PDA plus detector, a Finnigan LTQ linear ion trap with ESI source, and a Waters C18 Nova-Pak column (3.9 \times 100 mm, particle size 4 μm). The column temperature was maintained at 30°C and equilibrated with 95% solvent A (water/formic acid 99.9:0.1 v/v) and 5% solvent B (methanol/formic acid 99.9:0.1 v/v) at a flow rate of 1 ml min^{-1} , with 10% going to the mass spectrometer. Compounds were eluted by linear gradient to 65% solvent B over 60 min. Compounds were detected by PDA from 240 to 400 nm and MS in positive and negative mode from 150 to 1,500 m/z with the following MS parameters: sheath gas 30 and auxiliary gas 15 (both arbitrary units), spray voltage -4.0 kV in negative and 4.8 kV in positive ionization mode, capillary temperature 320°C , capillary voltage -1.0 and 45 V , respectively, and tube lens voltage -68 and 110 V , respectively.

Aglycones were identified by direct comparison with relevant flavonoid standards, and glycosylated flavonoid compounds were identified by comparison of UV spectra and MS2 fragmentation

patterns with standard flavonoid compounds. Standards were purchased from Carbosynth Ltd. (Compton, Berkshire, UK).

Statistical Analysis of Biochemical Traits Measured

The replicated mean data of 222 PMiGAP genotypes for the two traits *viz.*, FRAP and DPPH expressed in mg/ml was subjected to one factor analysis of variance (ANOVA) using the statistical software OPSTAT to find out the significant differences among the PMiGAP lines for both FRAP and DPPH activity (Sheoran et al., 1998). The means along with critical difference and coefficient of variation values were calculated using the software OPSTAT (Sheoran et al., 1998). The correlations between phenolic compounds and antioxidant activities were statistically evaluated by two-tailed bivariate correlate analysis, and were indicated by Pearson's coefficient indices, and $p < 0.05$ was considered as statistically significant. Heritability in broad sense was calculated using OPSTAT software (Sheoran et al., 1998) as the ratio of genotypic variance (σ^2_g) to the phenotypic variance (σ^2_p) and was expressed in the form of percentage (Hanson et al., 1956). The multivariate principal component analysis was performed by the JMPv.8 software (SAS Institute, 2008) to compare the multivariate correlation for antioxidant activities and phenolic compounds.

Genotyping and Filtering of Single-Nucleotide Polymorphisms

A set of SNPs resulting from 222 pearl millet accessions used in this study were obtained from ICRISAT (Varshney et al., 2017). The SNPs were filtered for site coverage (90%) and minimum minor allele frequency (MAF) of 0.01 with only bi-allelic markers using the Tassel ver. 5.2.64 software. Other SNP filtering criteria used in this study were no SSR motifs, no InDel marker, only bi-allelic SNPs, minor allele frequency (MAF) 0.01, and SNP quality score ≥ 30 . Genetic variant annotation for SNPs and its effect on genes and protein were predicted using SnpEff (<https://pcingola.github.io/SnpEff/>). Low-quality SNPs and SNPs with missing data were removed using the protocols described in Iquira et al. (2015) and Kujur et al. (2015).

Population Structure and Linkage Disequilibrium Mapping Analysis

The population structure analysis was performed using STRUCTURE software as described by Sehgal et al. (2015), which is based on Bayesian model-based cluster analysis (Pritchard et al., 2000). This was used for the assessment of patterns of genetic structure in 222 lines of the PMiGAP samples. This method used these 222 individuals to infer the fraction of an individual accession's genetic ancestry that belongs to a population, for a given number of populations ($K3-K10$). The "correct" K from the Ln probability of data [Ln P(D)], the delta- K values were estimated, as per the procedure suggested by Evanno et al. (2005). Maximum peak of ΔK was considered as the true cluster number.

The LD between pairs of polymorphic loci were analyzed using the software package TASSEL 5.2.64 (Bradbury et al., 2007;

<http://www.maizegenetics.net/>) as per the instructions of the user manual. LD was estimated using the squared allele frequency correlations (r^2), which is a measurement of the correlation between a pair of variables (Hill and Robertson, 1968). Decay of LD with genetic distance was estimated by nonlinear regression (SPSS Version 10.0) following the methods of Remington et al. (2001). The expected decay of LD was modeled as per Weir and Hill (1986).

Genome-Wide Association Analysis

Marker-trait association (MTA) analysis was performed using TASSEL program, ver. 5.2.64 by two different models, generalized linear model (GLM) and mixed linear model (MLM) (Bradbury et al., 2007). The MLM model-based analysis essentially required the population structure (Q), kinship (K), and a total of 67 K high-quality SNPs from 222 pearl millet lines for generation of K-matrix (Yu et al., 2006). The K-matrix was generated using the default parameters by choosing "Centered_IBS" method to obtain a better estimate of additive genetic variance. The default settings of the program were used for filtering marker data for minimum genotype count and minor allele frequency. Furthermore, no compression option in combination with P3D' for variance component estimation was adapted during MLM-based association analysis. The experiment-wise p -value provided a test of significance that corresponded to the experiment-wise error and was used to make decisions about the significantly associated markers. The markers were filtered at $-\log_{10} p\text{-value} \geq 3.0$ for considering as significant marker trait associations. The phenotypic variation (R^2) explained in percentage was also recorded for each individual marker. MTA analyses were also performed with an additional statistical method using the principal component analysis (PCA) matrix using the GLM (Yu et al., 2006). A permutation test using 1,000 permutations was allowed to correct the p -value for multiple comparisons. Q-Q plots and Manhattan plots were generated using the R package qqman (<https://cran.r-project.org/web/packages/qqman/index.html>; Turner, 2014).

Identification of Probable Candidate Genes

Significant SNPs from the GWAS were located by position and chromosome number on the *Pennisetum glaucum* reference genome (<ftp://cegresources.icrisat.org/>) using the CLC Genomics Workbench v.6.5 (CLC Bio, Aarhus, Denmark). Upstream and downstream regions surrounding each significantly associated SNP were searched in order to find and propose possible candidate gene. Furthermore, BLAST search of the NCBI database was conducted on each region of interest against the NCBI nr database.

RESULTS

Antioxidant Capacity Analysis by 1,1-Diphenyl-2-Picrylhydrazyl Free Radical Scavenging Capacity

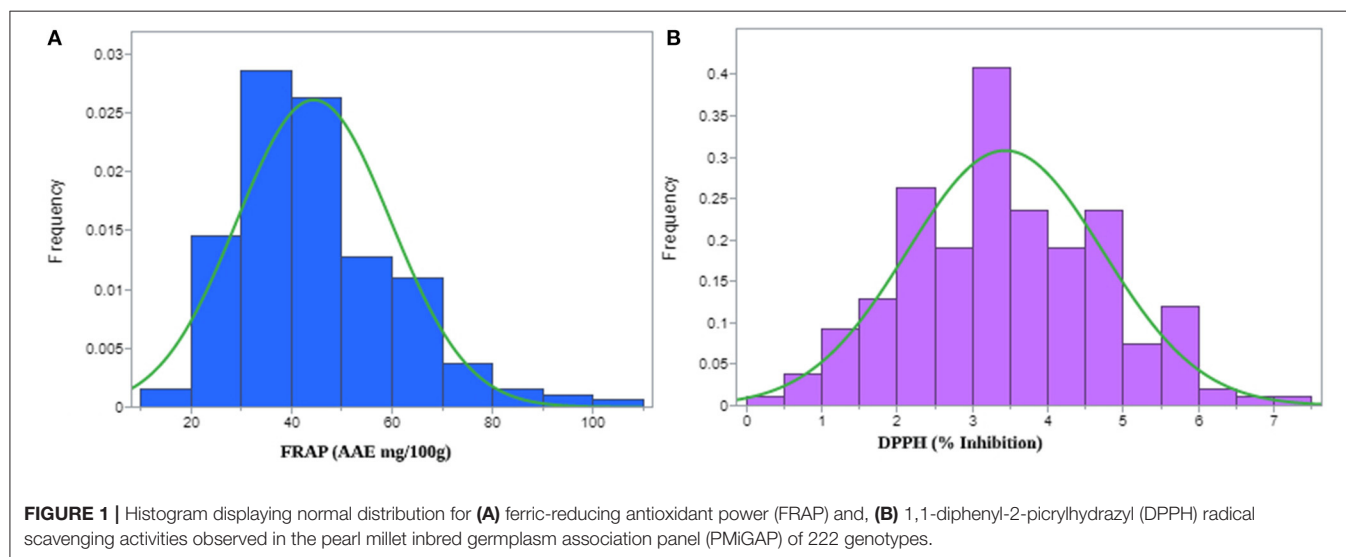
DPPH free radical scavenging activities were assessed for the pearl millet flour extracts. Table 1 displays the significant differences in DPPH activity among the 222 genotypes

TABLE 1 | Analysis of variance (ANOVA) for phenotypic antioxidant traits [1,1-diphenyl-2-picrylhydrazyl (DPPH) and ferric reducing antioxidant power (FRAP)] measured in triplicate for 222 pearl millet genotypes of the pearl millet inbred germplasm association panel (PMiGAP) (one factor).

Source of variation	Degree of freedom	Mean squares		F-calculated	
		FRAP (AAE ^a mg/100 g)	DPPH (% inhibition)	FRAP (AAE mg/100 g)	DPPH (% inhibition)
Genotypes	221	1,933.11	1,281.56	3,152.32**	3,758.92**
Error	442	0.61	0.34	-	-

** $p = 0.05$.

^aAAE, ascorbic acid equivalents.



of the PMiGAP. All the millet extracts scavenged DPPH radicals in a dose-dependent manner. DPPH scavenging activity (% inhibition) of the extracts ranged from 2.32% (in IP4965 genotype) to 112.45% (in IP10539 genotype) (Supplementary Table 3). Overall average DPPH scavenging activity of the extracts among pearl millet cultivars was found to be 53.75% (Figure 1A). These results indicated that extracts of pearl millet contain various phenolic antioxidants with the ability of donating hydrogen and scavenging free radicals. The heritability (measured as repeatability across three runs) was 99.92% for DPPH trait-regulatory factor and the phenotypic coefficient of variations for DPPH was 38.46.

Ferric-Reducing Antioxidant Power Assay

In this study, FRAP was calculated using the equation $Y = 0.2453x$, where x is the OD of the sample. The genotypic differences between the 222 genotypes were significant for FRAP activity (Table 1). Overall average FRAP activity of the extracts in pearl millet genotypes was 73.74 (AAE mg/100 g) (Figure 1B). The FRAP activity ranged from 21.68 mg/100 g (IP9446 genotype) to 179.66 mg/100 g (IP10579 genotype) of ascorbic acid (Supplementary Table 3). The highest ferric reducing ability of the extracts was found in IP7967 (150.57 mg/100g) and IP9824 (149.27 mg/100g). Higher antioxidant activity in these lines may be related to a higher content of phenolic compounds. The heritability of this extract (measured

as repeatability across three runs) for FRAP was calculated, and it was found to be very high (99.91%). The phenotypic coefficient of variations for FRAP was recorded as 34.44.

High-Performance Liquid Chromatography With Online Photodiode Array Detection and Electrospray Ionization-Ion Trap Tandem Mass Spectrometry Analysis

The total phenolics and flavonoid content that resulted from the HPLC-PDA-ESI/MSⁿ is presented in Figure 2, which shows the mass spectral data showing identified peaks in the representative chromatograms of the phenolic fractions of the selected millet grains. Phenolic compounds identified in millets were mainly of the flavonoids class, although spermidine hydroxycinnamate conjugates were also detected. Apigenin glycosides were the most prevalent flavones in the germplasm samples, and these included apigenin-8-C-glucoside (AH; Vitexin), apigenin-C-pentoside-C-pentoside (ADP), and apigenin-O-hexoside-C-hexoside (ADH). Luteolin-glycosides were also relatively abundant with luteolin-C-O-dihexosidedihex (LDH), and its caffeic acid conjugate was detected in extracts. The phenolamide, dicaffeoyl spermidine, was also observed in the sample extracts. An interesting pattern is seen in Figure 3, which demonstrates that apigenin is the predominant flavone in pearl millet samples with high levels of phenols, with the exception of Tift186, while for

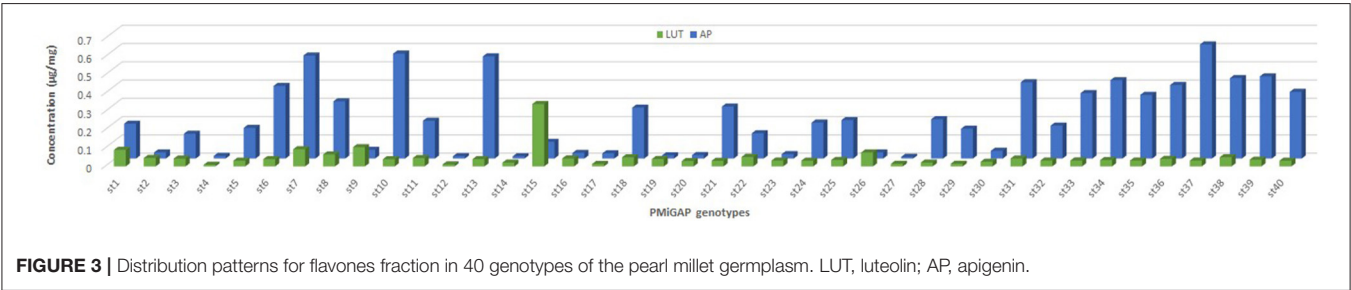
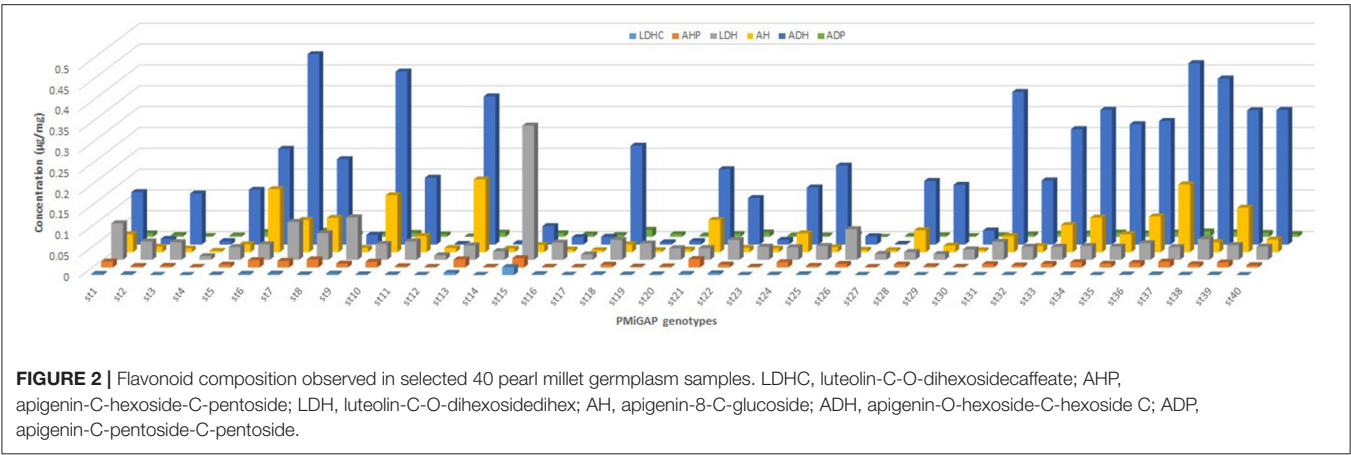


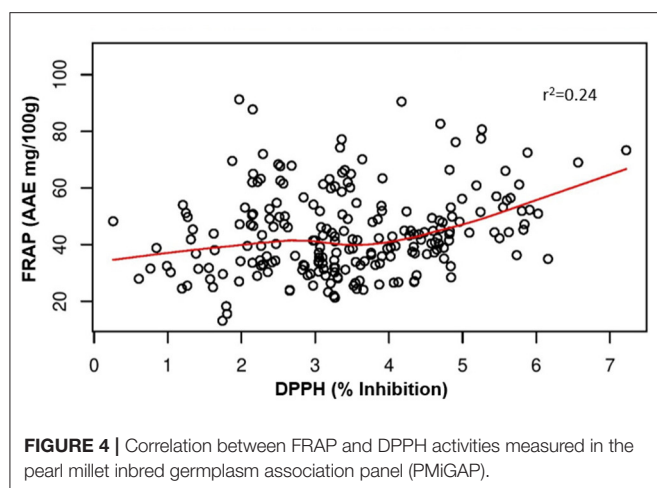
TABLE 2 | Correlation studies between phenolic compounds and antioxidant activities observed in selected 40 entries of the pearl millet germplasm.

	DPPH	FRAP	LHDC	AHP	LDH	AH	ADH	ADP	AP	LUT
DPPH	1	0.636*	0.572	0.880**	0.657	0.296	−0.491**	−0.420**	−0.392*	0.880**
FRAP		1	0.966**	0.792	0.965**	0.778	0.408	0.434	0.493	0.792
LHDC			1	0.546**	0.920**	0.074	−0.080	0.190	0.272	0.546**
AHP				1	0.501**	0.718**	0.540**	0.430**	0.762**	1.000**
LDH					1	−0.052	−0.052	0.212	0.286	0.501**
AH						1	0.724**	0.371*	0.788**	0.718**
ADH							1	0.653**	0.926**	0.540**
ADP								1	0.662**	0.430**
AP									1	0.762**
LUT										1

**Correlation is significant at the 0.01 level (two tailed).
*Correlation is significant at the 0.05 level (two tailed).
LDHC, luteolin-C-O-dihexosidecaffeate; AHP, apigenin-C-hexoside-C-pentoside; LDH, luteolin-C-O-dihexosidedihex; AH, apigenin-8-C-glucoside; ADH, apigenin-O-hexoside-C-hexoside C; ADP, apigenin-C-pentoside-C-pentoside; AP, apigenin; LUT, luteolin.

samples with lower total levels; apigenin is generally the major flavone core. The results also suggest that apigenin levels are far more variable than luteolin content. In millet phenolic extracts, several compounds from different flavonoid groups were detected that were either positively or tentatively identified as flavan-3-ol (monomers and dimers), flavonols, and their glycosides, and flavones, although in the current study, only flavone glycosides were found to be present in quantifiable levels.

Correlation Studies Between Antioxidant Activities and Phenolic Compounds
Polyphenols showed a positive and significant correlation with antioxidant activity (Table 2) as evident from the Pearson’s coefficients between phenolics and different antioxidant activities. DPPH showed highly significant positive correlation with AHP, luteolin (LUT) ($r^2 = 0.88^*$), and FRAP ($r^2 = 0.64^{**}$) indicating a strong association between phenolic compounds and antioxidant activity. FRAP also showed significant and



positive correlation with LHDC ($r^2 = 0.97^{**}$), and (LDH) ($r^2 = 0.96^{**}$). THL has also shown positive correlation with LDH ($r^2 = 0.92^{**}$) (Table 2). The results also clearly demonstrate a stronger correlation between total luteolin content and antioxidant activity ($r^2 = 0.88$ and 0.79 with DPPH and FRAP assays, respectively) compared with total apigenin content ($r^2 = -0.39$ and 0.49 with DPPH and FRAP assays, respectively) (Table 2). Multivariate principal component analysis (PCA), displayed the consistent dispersion pattern for each trait in relation to the first two principal coordinates. Based on measured eigenvalue, the first principal coordinate explained 39.2% of the total variation; the second coordinate could explain only 29.9%. The first five coordinates could explain a total of 96.01% variation. Thus, the first two components of the PCA explain a total of 69.07% of the variability among antioxidant activities and phenolic compounds (Supplementary Figure 1). The pairwise comparisons of DPPH and FRAP showed a weak correlation between the DPPH and FRAP with the r^2 value as 0.24 (Figure 4). This indicates that the radical scavenging activity measured by DPPH does not correspond directly with the reducing antioxidant activity measured by FRAP.

Identification and Distribution of Single-Nucleotide Polymorphisms

A total of 67,979 high-quality SNPs distributed all over the seven chromosomes of pearl millet genome were identified (Supplementary Figure 2). In general, one SNP change was observed at every 23,015 bases in the genome. The average nucleotide change rate was 43.35 SNPs per 1,000 bases of the genome. The maximum SNP change rate was observed on chromosome 5, which was 46.74 SNPs per Mb region of the genome and minimum on chromosome 4 (38.04). Structural annotation of 67,979 SNPs revealed the presence of 4,027 (~1.3%) SNPs in exonic regions, followed by 57,251 (19.4%) in intergenic regions. A total of 2,177 SNPs (0.73%) showed non-synonymous types of modification, whereas 1,831 (0.62%) were synonymous SNPs. A total of 9,605 (3.25%) SNPs were underlying from intergenic region,

and 294 (0.003%) SNPs were present in 5' UTR regions (Supplementary Figure 3).

Population Structure and Linkage Disequilibrium Analysis

The clustering program has estimated the membership probability (Q-matrix) of each PMiGAP accession to combine into a number of hypothetical subpopulations (K3–K10), and ΔK value was generated for subsequent runs. The optimum K at run 6 showed maximum peak during cluster analysis. In this way, Bayesian model-based cluster analysis revealed that the 222 individual genotypes were clustered into six groups ($K = 6$) (Supplementary Figure 4). These individuals were further classified into the ones with “pure” ancestry (where >50% of their inferred ancestry was derived from only one of the clusters) and “mixed ancestry” or “admixtures” (where <50% of inferred ancestry was derived from more than one cluster). The majority of the accessions belonged to the “mixed” ancestry, and the remaining 80 individuals from all six clusters were of “pure” ancestry.

The extent of LD was assessed among 67,979 pairs of loci (Supplementary Figure 5). Across all accessions, 2.08% of the total marker pairs showed a significant level of LD ($P < 0.01$). The average of r^2 for all pairs was 0.108, which may be attributable to low levels of markers. Counts for individual r^2 values were also analyzed by Varshney et al. (2017), where the r^2 threshold was set as 0.2, and rapid LD decay of <0.5 kb in PMiGAP lines (84–444 bp) was observed.

Markers Associated With 1,1-Diphenyl-2-Picrylhydrazyl and Ferric-Reducing Antioxidant Power

Marker trait association (MTA) analysis was performed for DPPH and FRAP activity-related trait using TASSEL. The MLM model-based association mapping showed 65 SNP markers associated with DPPH, whereas 153 SNPs showed association with FRAP trait at a p -value ≤ 0.001 (Supplementary Tables 4, 5). The MTA test resulted from the MLM model were further visualized into Manhattan plots in which each SNP was plotted against their chromosomal positions and the observed p -values (on a $-\log_{10}$ scale) to remark the highly significant SNP markers associated with a trait. Thus, higher stringent threshold values were considered to filter out the highly associated markers with a specific trait $-\log_{10}$ (p -value = 3.0) to minimize the effects of moderate size of the association panel and background noise of the Manhattan plots. Twenty-one SNP markers were found to be highly associated with the DPPH trait at p -value = 0.001, and Manhattan plot visualization revealed the high $-\log_{10}$ p -value ranging from 3.0 to 3.76 (Figure 5A). The percentage of phenotypic variation ranged from 4.8 to 10.4% for the DPPH trait. These markers were found to be distributed on chromosomes 1, 2, 3, and 6 for DPPH, which had the highest r^2 value. Out of 153, 80 SNP markers exhibited strongest associations with FRAP activity based on the MLM model at the p -value of 0.0001. Interestingly, 23 SNPs showed strongest associations with FRAP based on the MLM

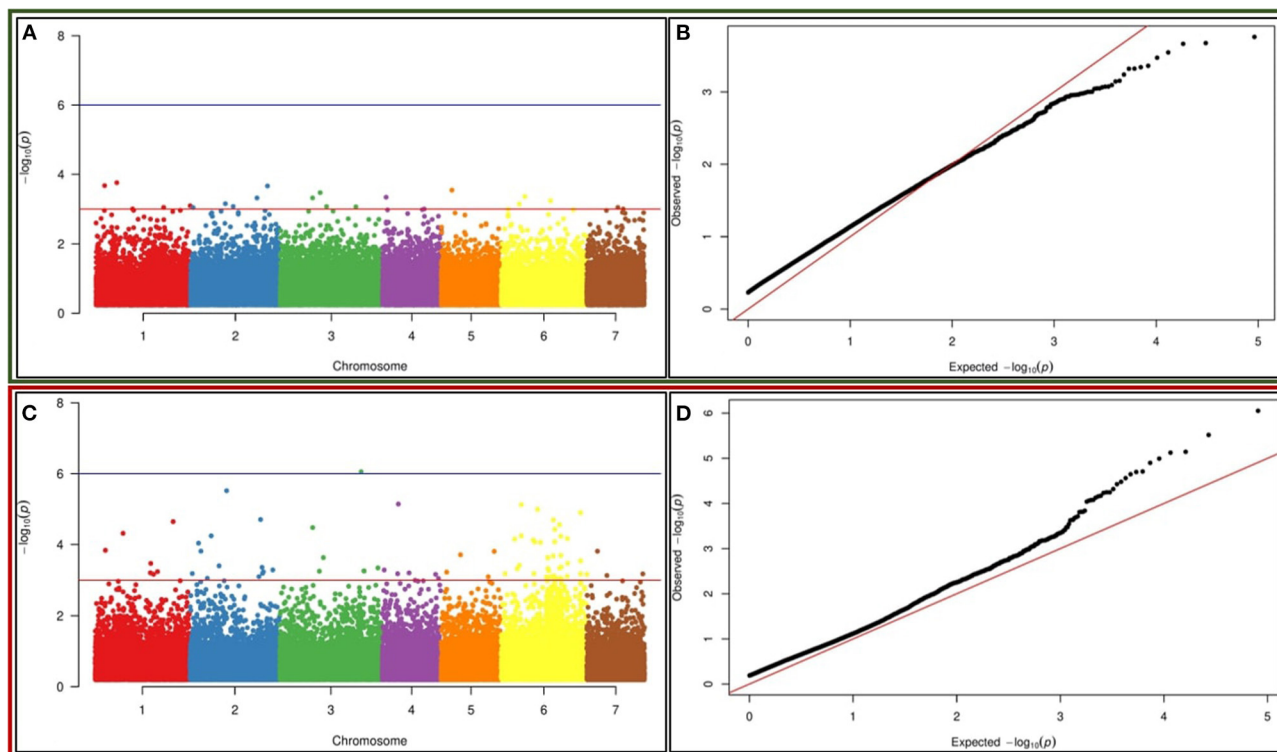


FIGURE 5 | Genome-wide association studies (GWAS)-based Manhattan plots built in the TASSEL v5.2.64 environment exhibiting significant p -values measured by mixed linear model (MLM) model for **(A)** DPPH **(C)** FRAP activity using 67 K SNPs in pearl millet. The x-axis illustrates the relative density of *Pennisetum glaucum* reference genome-based SNPs physically mapped on seven chromosomes. The y-axis displays the $-\log_{10} p$ -value for the significant association of SNP loci for DPPH the trait. Quantile–quantile plot for **(B)** DPPH and **(D)** FRAP activity using the MLM model, built in the TASSEL v5.2.64 environment. The x-axis displayed the expected $-\log_{10} p$ -value and y-axis represented the observed $-\log_{10} p$ -value.

model at the range of p -values = $9.5731\text{E}-04$ to $3.6325\text{E}-08$ (**Figure 5C**). The strongest association of these SNPs for FRAP trait was also explained by the Manhattan plot and the $-\log_{10} p$ -value ranges from 3.7–6.0. Maximum number of markers were distributed on chromosome 6 for this trait followed by chromosomes 1 and 2. Interestingly, the phenotypic variance ranged from 4.7 to 14.92% for these markers, and they were present on chromosomes 1, 2, and 6 for FRAP, which had the highest r^2 value. QQ (quantile–quantile) plots displayed linear distribution when plotted against the observed and expected distribution of p -values for both the traits (DPPH and FRAP) (**Figures 5B,D**).

A GLM model-based association mapping approach was also applied to identify the associated markers for comparison purposes. It found that 122 SNPs exhibited close association with DPPH and 388 SNP markers with FRAP at p -value ≤ 0.001 , but only common markers present in MLM-based analysis have been considered for further downstream analysis (**Supplementary Tables 4, 5**). After filtering, four SNP markers were found to be highly associated with DPPH at $-\log_{10} p$ -value ranging from 4.0 to 4.4 (**Supplementary Figure 6**). The phenotypic variance associated with them ranged between 4.4 and 10.4%, and they were distributed on chromosomes 1 and 6. Similarly, 60 SNP markers displayed strong association with

FRAP at a very high confidence statistical power. The Manhattan plot showed that the $-\log_{10} p$ -value scores in all cases were very high (3.0–7.4) (**Supplementary Figure 8**). QQ (quantile–quantile) plots displayed a linear distribution when plotted against the observed and expected distribution of p -values for both the traits (DPPH and FRAP) (**Supplementary Figures 7, 9**).

Identification of Candidate Genes

Furthermore, we investigated the candidate gene falling in the association region, which showed strong association with the trait of interest. The significant SNPs identified as a result of the GWAS were located by position on their respective chromosome of *Pennisetum glaucum* by reference genome assembly (https://cegresources.icrisat.org/data_public/PearlMillet_Genome/). A total of 436 gene surrounding each SNP was identified using the GFF file of the reference genome. A BLAST search against the NCBI database using Blast2go revealed a large number of hits, representing a vast number of candidate genes of which the functions included defense against biotic/abiotic stress factors, growth, development, and regulation of bioactive metabolites (**Supplementary Figure 10**). Out of 436, 18 genes, showed similarity with antioxidant biosynthetic pathway-related genes (**Table 3**). Functional annotation revealed that these

TABLE 3 | List of candidate genes residing around SNP markers found to be associated with DPPH and FRAP using generalized linear model (GLM) and mixed linear model (MLM) analysis on a collection of 222 individuals of the PMiGAP.

Associated SNP	Candidate gene	Chr	Start	End	Strand	Functional annotation
S3_254044945	Pgl_GLEAN_10001315	3	254,071,667	254,073,151	+	Flavanone 7-O-beta—glycosyltransferase
S6_180998971	Pgl_GLEAN_10005798	6	180,991,308	180,999,267	+	ATP binding cassette
S5_130817729	Pgl_GLEAN_10014356	5	130,814,880	130,815,485	-	Pathogenesis-related protein 1-like
S3_253451484	Pgl_GLEAN_10018369	3	253,448,279	253,448,695	-	NAC transcription factor
S2_34036202	Pgl_GLEAN_10025805	2	34,004,192	34,006,162	+	Anthocyanin 5-(6'''-hydroxycinnamoyltransferase)
S6_63963146	Pgl_GLEAN_10037383	6	63,946,271	63,948,373	+	Flavone 7-O-beta—glycosyltransferase
S3_230848548	Pgl_GLEAN_10010555	3	230,844,116	230,845,301	+	NAC domain containing protein
S1_42630463	Pgl_GLEAN_10009255	1	42,655,782	42,657,214	+	Flavanone 7-O-beta—glycoside
S7_131386682	Pgl_GLEAN_10024630	7	131,376,401	131,382,239	-	Kinase family protein
S5_153744160	Pgl_GLEAN_10009510	5	153,742,622	153,748,266	-	ATP binding cassette
S4_135618402	Pgl_GLEAN_10019136	4	135,631,750	135,633,026	+	Flanoid O-methyltransferase-like protein
S2_22074367	Pgl_GLEAN_10010945	2	21,992,330	21,994,014	-	Agmatine coumaroyltransferase-2
S6_84647706	Pgl_GLEAN_10019153	6	84,447,652	84,449,835	-	Putative DUF594 domain containing protein
S6_162685647	Pgl_GLEAN_10026569	6	163,032,154	163,035,976	+	DUF3755 family protein
S2_4439224	Pgl_GLEAN_10027118	2	4,490,408	4,491,589	+	Kinase family protein
S6_151551492	Pgl_GLEAN_10028864	6	151,435,588	151,435,863	-	Purple acid phosphatase 15
S6_146769180	Pgl_GLEAN_10037404	6	146,718,522	146,720,598	+	GDSL esterase/lipase
S4_136907468	Pgl_GLEAN_10012466	4	136,793,025	136,800,340	-	Glutathione S-transferase

genes contained molecular active sites involved in regulating antioxidant activities. For example, GDSL esterase/lipase (Pgl_GLEAN_10037404) was found to be involved in hydrolase activity and especially acting on ester bonds to promote the sequestration and accumulation of carotenoids. This suggests that these genes may be acting as co-factors in these pathways, or they may regulate genes through various chemical pathways to promote the accumulation of antioxidant compounds. Furthermore, some candidate genes were found to be directly associated with antioxidant biosynthetic pathways across all SNP datasets including UDP-glycosyltransferase 73D1-like (Pgl_GLEAN_10001315), agmatine coumaroyltransferase-2 (Pgl_GLEAN_10010945), anthocyanin 5-(6'''-hydroxycinnamoyltransferase) (Pgl_GLEAN_10025805), GDSL esterase/lipase (Pgl_GLEAN_10037404), purple acid phosphatase 15 (Pgl_GLEAN_10028864), DUF3755 family protein (Pgl_GLEAN_10026569), and ATP binding cassette (Pgl_GLEAN_10005798), which were deemed to be the most significant ($P = 6.84E-06$). These candidate genes will be taken forward for further verification by haplotyping and characterization of their functions in future studies.

DISCUSSION

Phytochemical antioxidants have numerous nutritional benefits, especially phenolic compounds, which are an important group of secondary metabolites with bioactive properties (e.g., hydroxycinnamates and flavonoids) and play a significant role in plants and human health conditions such as cancer, diabetes, and heart disease (Chandrasekara and Shahidi, 2011; Ofori et al., 2020). Pearl millet [*Penisetum glaucum* (L) R. Br.] is widely

cultivated as a dietary staple food in the arid and semi-arid regions of the world, particularly in India and Africa and known to be a good source of natural antioxidants. In the present study, we explored the antioxidant activity in the pearl millet germplasm (PMiGAP) set of 222 genotypes, by screening it for DPPH and FRAP assays. These assays are widely applied in numerous studies to investigate the free radical scavenging ability and for measuring antioxidant activity of natural extracts [Alam et al. (2013), López et al. (2014), and Berwal et al. (2016)]. The phenotypic evaluation and comparison of antioxidant activity in PMiGAP revealed distinct differences among pearl millet lines for antioxidant activities. The majority of pearl millet accessions were found to have higher FRAP activity. On an average, DPPH scavenging activity of 53.8% was observed among the pearl millet germplasm studied in this report. A similar trend of antioxidant capacity was also reported in other studies in pearl millet that demonstrated that antioxidant potential of pearl millet was greatly influenced by the cultivar (Fukumoto and Mazza, 2000; Pushparaj and Urooj, 2014; Kalteh et al., 2018). In five bran extracts of pearl millet, substantial levels of differences in total phenolics, flavonoids, and DPPH radical scavenging activities were reported (Iqbal et al., 2007). On a similar trend, Berwal et al. (2016) also reported a study on the antioxidant potential of 92 pearl millet genotypes by studying DPPH and ABTS radical assays.

Phenolic compounds such as phenolic acids and polyphenolic flavonoids are known to scavenge various free radicals by two different mechanisms, either by electron transfer or by hydrogen atom transfer, and protect cells and tissues of the body from oxidative stress, and thereby protecting humans from various diseases (Scalbert et al., 2005; Seifried et al., 2007; Valko et al., 2007). Flavones such as luteolin, apigenin, naringenin, and

their *o*-methyl derivatives are reported to have anti-cancerous properties (de Morais et al., 2017; Salazar-Lopez et al., 2018). Interestingly, in our study, DPPH showed a positive correlation with AHP, luteolin (LUT), suggesting higher phenolic compound accumulation and increased antioxidant activities. A similar trend of correlation with LHDC and LDH was also observed for FRAP activities. The similar observations were also recorded by Awika et al. (2004), Dykes et al. (2005), and Shen et al. (2018) where ADH showed positive correlation with AP ($r^2 = 93^{**}$). This is consistent with the theory that the dihydroxy functionality present in the luteolin molecule increases, reducing power in comparison with flavonoids with a single hydroxy group on the B ring as found with apigenin (Leopoldini et al., 2004).

Association mapping has been successfully performed in a number of crops including wheat, rice, pearl millet, maize, and cotton (Srivastava et al., 2020). In pearl millet, genetic diversity of PMiGAP panel and GWAS has been combined recently for several agronomic traits and reported recently (Varshney et al., 2017). However, association analysis of antioxidant trait has not been reported in pearl millet so far and is being reported for the first time *via* this study. A total of 222 genotypes studied in this report were drawn from the PMiGAP, which represented global diversity for antioxidant-related traits in pearl millet. The origin of each of these genotypes is provided in **Supplementary Table 1**, confirming that they are highly variable and were collected from different parts of the world. Population structure analysis for these accessions revealed that these individuals were grouped into six clusters in the panel. Interestingly, our results corroborate with earlier studies; for example, Sehgal et al. (2015) reported that there were six subpopulations within these PMiGAP by using 345 entries of the PMiGAP and 37 SSR markers. Similarly, population genomic analysis was carried out by Kanfany et al. (2020) and reported that the pearl millet inbred lines derived from diverse geographic and agroecological features possessed five subgroups mostly following pedigree differences with different levels of admixture. They have also stated that high genetic diversity prevails in pearl millet, which is very useful in defining heterotic groups for hybrid breeding, trait mapping, and holds promise for improving pearl millet for yield and nutritional quality. Six sub-groupings of pearl millet were also observed by Serba et al. (2019) using 398 accessions and 82,112 SNPs.

In the present study, we report the significantly associated MTAs for the antioxidant-related traits using ~67 K SNPs. Thus, a total of 24 SNPs strongly associated with the antioxidant-related traits (qualified at $-\log_{10} p\text{-value} \geq 3.0$) were further searched for LD status and for candidate gene searching. GWAS of these was performed by using widely used statistical models GLM and MLM (Yu et al., 2006; Price et al., 2010). In association analysis, the GLM procedure takes into account the phenotypic and genotype data combined with population structure data, while MLM considers both population structure and kinship data (familial relatedness), and therefore considered a stronger method, which avoids the chances of finding spurious associations (Zhao et al., 2007). However, in MLM, sometimes, over-compensation with both Q and K may also lead to false negatives (type II error; Zhao et al., 2007, 2011; Anuradha et al., 2017). Therefore, to overcome such difficulties, in this

study, both the statistical models (MLM and GLM) were applied for antioxidant association, and only the common significantly associated markers were considered to be more reliable. Like us, He et al. (2019) had also relied on common MTAs using GLM or MLM for target traits that have a normal distribution.

A total of 436 genes were identified around each SNP found to be associated with DPPH and FRAP activity. Eighteen candidate genes including UDP-glycosyltransferase, agmatine coumaroyltransferase-2, anthocyanin 5-(6''-hydroxycinnamoyltransferase), GDSL esterase/lipase, purple acid phosphatase 15, DUF3755 family protein, ATP binding cassette, and putative DUF594 domain-containing protein were found to be directly associated with antioxidant biosynthetic pathways across all selected SNP datasets. These candidates were found in close vicinity of the SNPs identified in this study but would require further validation in the future before they can be deployed in breeding programs. For validation, one can use candidate gene-based association mapping using large populations, or functional characterization through RNAi, VIGS, etc. Identification of desirable alleles of these MTAs will further confirm their roles in antioxidant synthesis and ensure their efficient utilization in crop improvement programs. Similar approaches as we report here were also adopted by Luo et al. (2020) for association analysis with variation in tocopherol and tocotrienol content, and these authors found that out of 47 associated SNPs, seven SNP markers were mapped in LD blocks ($r^2 > 0.6$), while 185 of the other SNPs were not mapped in LD blocks. Thus, they identified 88 genes, which were predicted to be transcription factors, including NAC domain transcript factors, and Myb domain transcripts. However, they selected EgHGGT (homogentisate geranylgeranyl transferase) for further analysis, which is involved in biosynthesis of tocotrienols and had higher expression levels in the mesocarp compared with other tissues. Antioxidant-associated candidate gene identification has previously been attempted by many researchers in maize and other cereal crops. In maize, association analysis was performed by Li et al. (2012) to identify SNP markers associated with phenotypic variation of vitamin E and found that an indel 85 kb upstream of the ZmVTE4 gene was detected and found to play a role in the expression of the ZmVTE4 gene. Similarly, association analysis was applied by Diepenbrock et al. (2017) to identify two chlorophyll biosynthetic enzyme genes, which were responsible for major variation in vitamin E content in maize, and they observed that HGGT and prephenate dehydratase had positive roles in phenotypic variation in vitamin E content. GWAS analysis was also performed by Varshney et al. (2017) in pearl millet using 3,117,056 SNPs on 20 agro-morphological traits, and strong association of the markers for GNP (grain number per panicle) trait was observed at chromosomes 1 and 5. Association mapping and significant allelic association for grain iron and zinc were also demonstrated by Anuradha et al. (2017) in pearl millet. Similarly, association mapping analysis was also performed by Saïdou et al. (2014) and Sehgal et al. (2015) for many agronomically important traits in pearl millet.

In conclusion, the present study explored the genetic architecture of antioxidant content in pearl millet for the first time using LD-based GWAS exploiting historical recombination

in a natural germplasm population. The population used showed a wide range of variability for traits studied. Furthermore, resequencing data and quality phenotyping of DPPH and FRAP led us to the identification of highly significant SNP markers associated with antioxidant content. Such associations provided better insights into the genetic architecture of this very important trait in pearl millet. Candidate gene analysis in the significantly associated SNP region has identified potential candidates, which on further validation, will provide a great resource to select these important traits in pearl millet breeding programs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

RSY conceived and supervised the completed study. JT, DY, PG, AW, and RY contributed to the phenotyping experiments. CY analyzed the results and wrote the manuscript. RKS and RBS developed the HapMap used for GWAS in this study. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

RSY would like to acknowledge funding received from the Innovate UK projects 102576 and 105672. CY acknowledges the Royal Society for providing the Newton International Fellowship. JT and DY acknowledge the Indian Council of Agricultural Research, Government of India, for the financial support and the CCSHAU for granting them leave for foreign training at the Aberystwyth University under the NAHEP—Institutional Development Plan (IDP).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.599649/full#supplementary-material>

REFERENCES

- Adom, K. K., and Liu, R. H. (2002). Antioxidant activity of grains. *J. Agric. Food Chem.* 50, 6182–6187. doi: 10.1021/jf0205099
- Akpanika, G. A., Winters, A., Wilson, T., Ayoola, G. A., Adepoju-Bello, A. A., and Hauck, B. (2017). Polyphenols from *Allanblackia floribunda* seeds: identification, quantification and antioxidant activity. *Food Chem.* 222, 35–42. doi: 10.1016/j.foodchem.2016.12.002
- Alam, M. N., Bristi, N. J., and Rafiquzzaman, M. (2013). Review on *in vivo* and *in vitro* methods evaluation of antioxidant activity. *Saudi Pharm. J.* 21, 143–152. doi: 10.1016/j.jsps.2012.05.002
- Ali, M. A. M., El Tinay, A. H., and Abdalla, A. H. (2003). Effect of fermentation on the *in vitro* protein digestibility of pearl millet. *Food Chem.* 80, 51–54. doi: 10.1016/S0308-8146(02)00234-0
- Anuradha, N., Satyavathi, C. T., Bharadwaj, C., Nepolean, T., Sankar, S. M., Singh, S. P., et al. (2017). Deciphering genomic regions for high grain iron and zinc content using association mapping in pearl millet. *Front. Plant Sci.* 8:412. doi: 10.3389/fpls.2017.00412
- Awika, J. M., Rooney, L. W., and Waniska, R. D. (2004). Properties of 3-deoxyanthocyanins from sorghum. *J. Agric. Food Chem.* 52, 4388–4394. doi: 10.1021/jf049653f
- Benzie, I. F. F., and Strain, J. J. (1999). Ferric reducing/antioxidant power assay: direct measure of total antioxidant activity of biological fluids and

Supplementary Figure 1 | Principal Component analysis (PCA) depicting variation among phenotypic traits for antioxidant activities and phenolic compounds.

Supplementary Figure 2 | Single nucleotide polymorphism (SNP) density on seven chromosomes of pearl millet. The x-axis shows the interval distance in Mb. Window size to calculate SNP density 1 Mb.

Supplementary Figure 3 | Annotation of single-nucleotide polymorphisms (SNPs) against reference genome of pearl millet showing distribution of SNPs in intergenic, upstream in different genic regions, synonymous, and non-synonymous SNPs detected within the CDS region.

Supplementary Figure 4 | Determination of optimum K using Evanno's method identified during population structure analyses.

Supplementary Figure 5 | LD plot built in the TASSEL v5.2.38 environment zoomed in on regions of high LD on chromosomes. The squared correlation coefficient (r^2) values are denoted by a color scale from white (0.0) to red (1.0) in the upper triangle. The p -values ranging from non-significant (0.01; white) to highly significant (<0.0001 ; red) are shown in the lower triangle.

Supplementary Figure 6 | GWAS-based Manhattan plots built in the TASSEL v5.2.64 environment exhibiting significant p -values measured by GLM model for DPPH activity using 67 K SNPs in pearl millet. The x-axis represented by SNPs on seven chromosomes and y-axis exhibited the $-\log_{10} p$ -value for the significant association of SNP loci for both the trait DPPH.

Supplementary Figure 7 | Quantile-quantile plot of DPPH using the GLM model, built in the TASSEL v5.2.64 environment.

Supplementary Figure 8 | GWAS-based Manhattan plots built in the TASSEL v5.2.64 environment exhibiting significant p -values measured by GLM model for FRAP activity. The x-axis displayed the distribution of SNPs throughout all pearl millet chromosomes and y-axis represented the $-\log_{10} p$ -value for the significant association of SNP loci for FRAP.

Supplementary Figure 9 | Quantile-quantile plot representing distribution pattern for SNP marker associated with FRAP trait along with p -value analyzed through GLM model.

Supplementary Figure 10 | Functional annotation of biological processes of genes residing around traits (DPPH and FRAP) associated SNPs.

Supplementary Table 1 | Details of the 222 accessions of the PMiGAP genotypes having origin from around the world and used in this study.

Supplementary Table 2 | Sample details of the pearl millet genotypes used for HPLC analysis.

Supplementary Table 3 | Phenotyping values of 222 PMiGAP pearl millet genotypes represented by mean values for FRAP assay and DPPH activity.

Supplementary Table 4 | Significantly associated SNP markers with DPPH at $P < 0.001$ analyzed through MLM and GLM based statistical model on a collection of 222 individuals of the PMiGAP.

Supplementary Table 5 | List of markers associated with FRAP trait at $P < 0.001$ using MLM and GLM based methods on a collection of 222 individuals of the PMiGAP.

- modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. *Methods Enzymol. Elsevier* 5, 15–27. doi: 10.1016/S0076-6879(99)99005-5
- Berwal, M. K., Chugh, L. K., Goyal, P., and Kumar, R. (2016). Total antioxidant potential of pearl millet genotypes: inbreds and designated b-lines. *Indian J. Agric. Biochem.* 29, 201–204. doi: 10.5958/0974-4479.2016.00032.0
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Buckler, E. S., and Gore, M. (2007). An Arabidopsis haplotype map takes root. *Nat. Genet.* 39, 1056–1057. doi: 10.1038/ng0907-1056
- Casa, A. M., Pressoir, G., Brown, P. J., Mitchell, S. E., Rooney, W. L., Tuinstra, M. R., et al. (2008). Community resources and strategies for association mapping in Sorghum. *Crop. Sci.* 48, 30–40. doi: 10.2135/cropsci2007.02.0080
- Chandrasekara, A., and Shahidi, F. (2011). Antiproliferative potential and DNA scission inhibitory activity of phenolics from whole millet grains. *J. Funct. Foods* 3, 159–170. doi: 10.1016/j.jff.2011.03.008
- Choi, Y., Jeong, H. S., and Lee, J. (2007). Antioxidant activity of methanolic extracts from some grains consumed in Korea. *Food Chem.* 103, 130–138. doi: 10.1016/j.foodchem.2006.08.004
- de Moraes, C. L., Pinheiro, S. S., Martino, H. S. D., and Pinheiro-Sant'Ana, H. M. (2017). Sorghum (*Sorghum bicolor* L.): nutrients, bioactive compounds, and potential impact on human health. *Crit. Rev. Food Sci. Nutr.* 57, 372–390. doi: 10.1080/10408398.2014.887057
- Diepenbrock, C. H., Kandianis, C. B., Lipka, A. E., Magallanes, L. M., Vaillancourt, B., Gongora, C., et al. (2017). Novel loci underlie natural variation in vitamin E levels in maize grain. *Plant Cell* 29, 23734–22392. doi: 10.1105/tpc.17.00475
- Dykes, L., Rooney, L. W., Waniska, R. D., and Rooney, W. L. (2005). Phenolic compounds and antioxidant activity of sorghum grains of varying genotypes. *J. Agric. Food Chem.* 53, 6813–6818. doi: 10.1021/jf050419e
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- FAOSTAT database (2017). *FAOSTAT Database, Food and Agriculture Organization of the United Nations, Rome, Italy.*
- Flint-Garcia, S. A., Thuillet, A. C., Yu, J. M., Pressoir, G., Romero, S. M., Mitchell, S. E., et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064. doi: 10.1111/j.1365-313X.2005.02591.x
- Fukumoto, L. R., and Mazza, G. (2000). Assessing antioxidant and prooxidant activities of phenolic compounds. *J. Agric. Food Chem.* 48, 3597–3604. doi: 10.1021/jf000220w
- Hanson, G. H., Robinson, H. F., and Comstock, R. E. (1956). Biometrical studies of yield in segregating population of Korean Lespedeza. *Agro. J.* 48, 267–282. doi: 10.2134/agronj1956.00021962004800060008x
- Hauck, D. (2018). New natural synergistic antioxidant blend for product protection. *SOFW J.* 144, 1–12.
- He, Q., Xu, F., Min, M. H., Chu, S. H., Kim, K. W., and Park, Y. J. (2019). Genome-wide association study of vitamin E using genotyping by sequencing in sesame (*Sesamum indicum*). *Genes Genom.* 41, 1085–1093. doi: 10.1007/s13258-019-00837-3
- Hill, W. G., and Robertson, A. (1968). The effects of inbreeding at loci with heterozygote advantage. *Genetics* 60, 615–628. doi: 10.1093/genetics/60.3.615
- Ilesanmi, O. K. B., and Akinloye, O. A. (2013). Assessment of nutritional composition and antioxidant ability of pearl millet (*Pennisetum glaucum*). *Am. J. Res. Commun.* 1, 262–272.
- Iqbal, S., Bhanger, M. I., and Anwar, F. (2007). Antioxidant properties and components of bran extracts from selected wheat varieties commercially available in Pakistan. *LWT Food Sci. Technol.* 40, 361–367. doi: 10.1016/j.lwt.2005.10.001
- Iqura, E., Humira, S., and Francois, B. (2015). Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biol.* 15, 1–12. doi: 10.1186/s12870-014-0408-y
- Jin, L., Lu, Y., Xiao, P., Sun, M., Corke, H., and Bao, J. (2010). Genetic diversity and population structure of a diverse set of rice germplasm for association mapping. *Theor. Appl. Genet.* 121, 475–487.
- Kalteh, M., Alipour, Z. T., Ashraf, S., Marashi, A. M., and Falah, N. A. (2018). Effect of silica nanoparticles on basil (*Ocimum basilicum*) under salinity stress. *J. Chem. Heal. Risks* 4, 49–55.
- Kanfany, G., Serba, D. D., Rhodes, D., Amand, P. S., Bernardo, A., Gangashetty, P. I., et al. (2020). Genomic diversity in pearl millet inbred lines derived from landraces and improved varieties. *BMC Geno.* 21, 1–12. doi: 10.1186/s12864-020-06796-4
- Khan, I., Yousif, A. M., Johnson, S. K., and Gamlath, S. (2015). Acute effect of sorghum flour-containing pasta on plasma total polyphenols, antioxidant capacity and oxidative stress markers in healthy subjects: a randomised controlled trial. *Clin. Nutr.* 34, 415–421. doi: 10.1016/j.clnu.2014.08.005
- Kujur, A., Bajaj, D., Upadhyaya, H. D., Das, S., Ranjan, R., Shree, T., et al. (2015). A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci. Rep.* 5, 11166. doi: 10.1038/srep11166
- Leopoldini, M., Marino, T., Russo, N., and Toscano, M. (2004). Antioxidant properties of phenolic compounds: H-atom versus electron transfer mechanism. *J. Phys. Chem. A* 108, 4916–4922. doi: 10.1021/jp037247d
- Li, Q., Yang, X., Xu, S., Cai, Y., Zhang, D., Han, Y., et al. (2012). Genome-wide association studies identified three independent polymorphisms associated with a-tocopherol content in maize kernels. *PLoS ONE* 7:e36807. doi: 10.1371/journal.pone.0036807
- López, S. A., Olivas, R. M., Landaluze, J. S., and Cámara, C. (2014). Nanoparticles: a global vision. Characterization, separation, and quantification methods. Potential environmental and health impact. *Anal. Methods* 6, 38–56. doi: 10.1039/C3AY40517F
- Luo, T., Xia, W., Gong, S., Mason, A. S., Li, Z., Liu, R., et al. (2020). Identifying vitamin E biosynthesis genes in *Elaeis guineensis* by genome-wide association study. *J. Agric. Food Chem.* 68, 678–685. doi: 10.1021/acs.jafc.9b03832
- Maccaferri, M., Sanguineti, M. C., Natoli, V., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2006). A panel of elite accessions of durum wheat (*Triticum durum* Desf.) suitable for association mapping studies. *Plant Genet. Resour.* 4, 79–85. doi: 10.1079/PGR2006117
- Muthamilarasan, M., Dhaka, A., Yadav, R., and Prasad, M. (2016). Exploration of millet models for developing nutrient rich graminaceous crops. *Plant Sci.* 242, 89–97. doi: 10.1016/j.plantsci.2015.08.023
- Osofu, F. K., Elahi, F., Daliri, E. B.-M., Chelliah, R., Ham, H. J., Kim, J.-H., et al. (2020). Phenolic profile, antioxidant, and antidiabetic potential exerted by millet grain varieties. *Antioxidants* 9:254. doi: 10.3390/antiox9030254
- Pandey, K. B., and Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxid. Med. Cell Longev.* 2, 270–278. doi: 10.4161/oxim.2.5.9498
- Price, A. L., Zaitlen, N. A., Reich, D., and Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* 11, 459–463. doi: 10.1038/nrg2813
- Pritchard, J. K., Stephens, M., Rosenberg, N. A., and Donnelly, P. (2000). Association mapping in structured populations. *Am. J. Hum. Genet.* 67, 170–181. doi: 10.1086/302959
- Pushparaj, F. S., and Urooj, A. (2014). Antioxidant activity in two pearl millet (*Pennisetum typhoideum*) cultivars as influenced by processing. *Antioxidants* 3, 55–66. doi: 10.3390/antiox3010055
- Ragae, S., Abdel-Aal, E. S. M., and Noaman, M. (2006). Antioxidant activity and nutrient composition of selected cereals for food use. *Food Chem.* 98, 32–38. doi: 10.1016/j.foodchem.2005.04.039
- Ramya, R. A., Ahmed, L., Satyavathi, C. T., Rathore, A., Katiyar, P., Bhaskar Raj, A. G., et al. (2018). Towards defining heterotic gene pools using SSR markers in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 8:1934. doi: 10.3389/fpls.2017.01934
- Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., et al. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11479–11484. doi: 10.1073/pnas.201394398
- Saïdou, A. A., Cloutault, J., Couderc, M., Mariac, C., Devos, K. M., and Thuillet, A. C. (2014). Association mapping, patterns of linkage disequilibrium and selection in the vicinity of the PHYTOCHROME C gene in pearl millet. *Theor. Appl. Genet.* 127, 19–32. doi: 10.1007/s00122-013-2197-3
- Salazar-Lopez, N. J., Gonzalez, A. G., Rouzaud, S. O., and Robles, S. M. (2018). Technologies applied to sorghum (*Sorghum bicolor* L. Moench): changes in

- phenolic compounds and antioxidant capacity. *Food Sci. Technol.* 38, 369–382. doi: 10.1590/fst.16017
- Saleh, A. S. M., Zhang, Q., Chen, J., and Shen, Q. (2013). Millet grains: nutritional quality, processing, and potential health benefits. *Compr. Rev. Food Sci. Food Saf.* 12, 281–295. doi: 10.1111/1541-4337.12012
- Scalbert, A., Johnson, I. T., and Saltmarsh, M. (2005). Polyphenols: antioxidants and beyond. *Am. J. Clin. Nutr.* 81, 215S–217S. doi: 10.1093/ajcn/81.1.215S
- Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., and Taunk, J. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS ONE* 10, 1–28. doi: 10.1371/journal.pone.0122165
- Seifried, H. E., Anderson, D. E., Fisher, E. I., and Milner, J. A. (2007). A review of the interaction among dietary antioxidants and reactive oxygen species. *J Nutr Biochem* 18:567–579 doi: 10.1016/j.jnutbio.2006.10.007
- Serba, D. D., Muleta, K. T., Amand, P. S., Bernardo, A., Bai, G., Perumal, R., et al. (2019). Genetic diversity, population structure, and linkage disequilibrium of pearl millet. *Plant Genome* 12, 1–12. doi: 10.3835/plantgenome2018.11.0091
- Serea, C., and Barna, O. (2011). Phenolic content and antioxidant activity in oat. *Food Sci. Technol.* 12, 164–168.
- Shahidi, F. (2000). Antioxidants in food and food antioxidants. *Nahrung Food* 44, 158–163. doi: 10.1002/1521-3803(20000501)44:3<158::AID-FOOD158>3.0.CO;2-L
- Shahidi, F., and Chandrasekara, A. (2013). Millet grain phenolics and their role in disease risk reduction and health promotion: a review. *J. Funct. Foods* 5, 570–581. doi: 10.1016/j.jff.2013.02.004
- Shen, S., Huang, R., Li, C., Wu, W., Chen, H., Shi, J., et al. (2018). Phenolic compositions and antioxidant activities differ significantly among sorghum grains with different applications. *Molecules* 23:1203. doi: 10.3390/molecules23051203
- Sheoran, O. P., Tonk, D. S., Kaushik, L. S., Hasija, R. C., and Pannu, R. S. (1998). “Statistical software package for agricultural research workers,” in *Recent Adv. Inf. Theory Stat. Comput. Appl.*, eds D. S. Hooda, R. C. Hasija (Hisar: Dep Math Stat CCS HAU), 139–143.
- Singh, B., Sharma, H. K., and Sarkar, B. C. (2012). Optimization of extraction of antioxidants from wheat bran (*Triticum* spp.) using response surface methodology. *J. Food Sci. Technol.* 49, 294–308. doi: 10.1007/s13197-011-0276-5
- Srivastava, R., Singh, R. B., Vijay, L. P., Srikanth, B., Satyavathi, C. T., Yadav, R. S., et al. (2020). Genome-wide association studies and genomic selection in pearl millet: advances and prospects. *Front. Genet.* 10, 1389–1402. doi: 10.3389/fgene.2019.01389
- Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *BioRxiv*. doi: 10.1101/005165
- Upadhyaya, H. D., Reddy, K. N., and Sastry, D. V. S. R. (2008). “Regeneration guidelines: pearl millet,” in *Crop Specific Regeneration Guidelines [CD-ROM]*, eds M. E. Dulloo, I. Thormann, M. A. Jorge, and J. Hanson (Rome: CGIAR System-wide Genetic Resource Programme), 9.
- Valko, M., Leibfritz, D., Moncol, J., Cronin, M. T. D., Mazur, M., and Telser, J. (2007). Free radicals and antioxidants in normal physiological functions and human disease. *Int. J. Biochem. Cell Biol.* 39, 44–84. doi: 10.1016/j.biocel.2006.07.001
- Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976. doi: 10.1038/nbt.3943
- Weir, B. S., and Hill, W. G. (1986). Nonuniform recombination within the human beta-globin gene cluster. *Am. J. Hum. Genet.* 38, 776–781.
- Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Yu, J. M., and Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17, 1–6. doi: 10.1016/j.copbio.2006.02.003
- Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., et al. (2007). An Arabidopsis example of association mapping in structured samples. *PLoS Genet.* 3:e4. doi: 10.1371/journal.pgen.0030004
- Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2, 467–477. doi: 10.1038/ncomms1467
- Zhu, C. S., Gore, M., Buckler, E. S., and Yu, J. M. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1, 5–20. doi: 10.3835/plantgenome2008.02.0089
- Zieliński, H., and Kozłowska, H. (2000). Antioxidant activity and total phenolics in selected cereal grains and their different morphological fractions. *J. Agric Food Chem.* 48, 2008–2016. doi: 10.1021/jf990619o

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Yadav, Tokas, Yadav, Winters, Singh, Yadav, Gangashetty, Srivastava and Yadav. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Performance and Stability of Pearl Millet Varieties for Grain Yield and Micronutrients in Arid and Semi-Arid Regions of India

OPEN ACCESS

Edited by:

Marcello Mastrorilli,
Council for Agricultural and
Economics Research (CREA), Italy

Reviewed by:

Pangirayi Tongoono,
University of KwaZulu-Natal,
South Africa
John Derera,
International Institute of Tropical
Agriculture (IITA), Nigeria

*Correspondence:

P. Sanjana Reddy
sanjana@millets.res.in
Mahalingam Govindaraj
m.govindaraj@cgiar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 20 February 2021

Accepted: 30 April 2021

Published: 31 May 2021

Citation:

Sanjana Reddy P, Satyavathi CT,
Khandelwal V, Patil HT, Gupta PC,
Sharma LD, Mungra KD, Singh SP,
Narasimhulu R, Bhadarge HH,
Iyanar K, Tripathi MK, Yadav D,
Bhardwaj R, Talwar AM, Tiwari VK,
Kachole UG, Sravanti K, Shanthi
Priya M, Athoni BK, Anuradha N,
Govindaraj M, Nepolean T and
Tonapi VA (2021) Performance and
Stability of Pearl Millet Varieties for
Grain Yield and Micronutrients in Arid
and Semi-Arid Regions of India.
Front. Plant Sci. 12:670201.
doi: 10.3389/fpls.2021.670201

P. Sanjana Reddy^{1*}, C. Tara Satyavathi², Vikas Khandelwal², H. T. Patil³, P. C. Gupta⁴,
L. D. Sharma⁵, K. D. Mungra⁶, Sumer P. Singh⁷, R. Narasimhulu⁸, H. H. Bhadarge⁹,
K. Iyanar¹⁰, M. K. Tripathi¹¹, Devvart Yadav¹², Ruchika Bhardwaj¹³, A. M. Talwar¹⁴,
V. K. Tiwari¹⁵, U. G. Kachole¹⁶, K. Sravanti¹⁷, M. Shanthi Priya¹⁸, B. K. Athoni¹⁹,
N. Anuradha²⁰, Mahalingam Govindaraj^{21*}, T. Nepolean¹ and Vilas A. Tonapi¹

¹ Indian Institute of Millets Research, Hyderabad, India, ² All India Coordinated Research Project on Pearl Millet, Jodhpur, India, ³ Bajra Research Scheme, College of Agriculture, Mahatma Phule Krishi Vidyapeeth (MPKV), Dhule, India, ⁴ Agricultural Research Station, Swami Keshavanand Rajasthan Agriculture University (SKRAU), Bikaner, India, ⁵ Rajasthan Agricultural Research Institute, Sri Karan Narendra Agriculture University (SKNAU), Jaipur, India, ⁶ Pearl Millet Research Station, Junagadh Agricultural University (JAU), Jamnagar, India, ⁷ ICAR-Indian Agricultural Research Institute (IARI), New Delhi, India, ⁸ Agricultural Research Station, Acharya NG Ranga Agricultural University (ANGRAU), Ananthapuramu, India, ⁹ National Agricultural Research Project, Vasantrao Naik Marathwada Krishi Vidyapeeth (VNMKV), Aurangabad, India, ¹⁰ School of Genetics, Tamil Nadu Agricultural University (TNAU), Coimbatore, India, ¹¹ College of Agriculture, Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya (RVSKVV), Gwalior, India, ¹² Chaudhary Charan Singh Haryana Agricultural University (CCSHAU), Hisar, India, ¹³ College of Agriculture, Punjab Agricultural University (PAU), Ludhiana, India, ¹⁴ Agricultural Research Station, University of Agricultural Sciences (UAS), Raichur, India, ¹⁵ Zonal Agricultural Research Station, Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya (RVSKVV), Morena, India, ¹⁶ Agricultural Research Station, Mahatma Phule Krishi Vidyapeeth (MPKV), Niphad, India, ¹⁷ Regional Agricultural Research Station, Professor Jayashankar Telangana State Agricultural University (PJTSAU), Palem, India, ¹⁸ Agricultural Research Station, Acharya NG Ranga Agricultural University (ANGRAU), Tirupati, India, ¹⁹ Regional Agricultural Research Station, University of Agricultural Sciences (UAS), Vijayapura, India, ²⁰ Agricultural Research Station, Acharya NG Ranga Agricultural University (ANGRAU), Vizianagaram, India, ²¹ International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is grown under both arid and semi-arid conditions in India, where other cereals are hard to grow. Pearl millet cultivars, hybrids, and OPVs (open pollinated varieties) are tested and released by the All India Coordinated Research Project on Pearl Millet (AICRP-PM) across three zones (A₁, A, and B) that are classified based on rainfall pattern. Except in locations with extreme weather conditions, hybrids dominate pearl millet growing areas, which can be attributed to hybrid vigor and the active role of the private sector. The importance of OPVs cannot be ruled out, owing to wider adaptation, lower input cost, and timely seed availability to subsidiary farmers cultivating this crop. This study was conducted to scrutinize the presently used test locations for evaluation of pearl millet OPVs across India, identify the best OPVs across locations, and determine the variation in grain Fe and Zn contents across locations in these regions. Six varieties were evaluated across 20 locations in A₁ and A (pooled as A) and B zones along with three common checks and additional three zonal adapted checks in the respective zones during the 2019 rainy season. Recorded data on yield and quality traits were analyzed using genotype main effects and genotype × environment

interaction biplot method. The genotype × environment (G × E) interaction was found to be highly significant for all the grain yield and agronomic traits and for both micronutrients (iron and zinc). However, genotypic effect (G) was four (productive tillers) to 49 (grain Fe content) times that of G × E interaction effect for various traits across zones that show the flexibility of OPVs. Ananthapuramu is the ideal test site for selecting pearl millet cultivars effectively for adaptation across India, while Ananthapuramu, Perumallapalle, and Gurugram can also be used as initial testing locations. OPVs MP 599 and MP 600 are identified as ideal genotypes, because they showed higher grain and fodder yields and stability compared with other cultivars. Iron and zinc concentration showed highly significant positive correlation (across environment = 0.83; $p < 0.01$), indicating possibility of simultaneous effective selection for both traits. Three common checks were found to be significantly low yielders than the test entries or zonal checks in individual zones and across India, indicating the potential of genetic improvement through OPVs.

Keywords: representative, G × E, GGE-biplot, iron, zinc, grain yield, fodder yield

INTRODUCTION

Pearl millet (*Pennisetum glaucum* L.R.Br.) is cultivated in dry regions of arid and semi-arid tropics where no other cereal can be successfully grown. India is the largest producer of millets in the world, harvesting about 11 million tons per year, nearly 36% of the world's output. Pearl millet, which accounts for about two-thirds of millet production in India, is grown in the drier areas of the country, mainly in the states of Rajasthan, Maharashtra, Gujarat, Uttar Pradesh, and Haryana. In India, pearl millet is the fourth most widely cultivated food crop after rice, wheat, and maize. It occupies an area of 6.93 million ha with an average production of 8.61 million tons and productivity of $1,243 \text{ kg ha}^{-1}$ (Directorate of Millets Development, 2020). The cultivated area for pearl millet in India is divided into three main zones based on soils and rainfall patterns. The northwestern part of India, receiving <400 mm of annual rainfall, is classified as an A₁ zone. The northern and central parts of India, with sandy loam soils and receiving >400 mm of annual rainfall, are denoted as an A zone, and the peninsular region of India, receiving >400 mm of annual rainfall and bearing heavy soils, is broadly classified as a B zone (Rai et al., 2015). The arid tracts are grown with landraces/OPVs (open pollinated varieties) that are poor yielders. The progress achieved in pearl millet yields is attributed to the active role of the private sector in the dissemination of pearl millet hybrids in the productive zone of northern and central India rather than in the arid zone (Rai et al., 2015). On the other hand, the public sector could not record progress on par with that of the private sector. The active role of the private sector and the predominantly cross-pollinated nature of the crop have led to the rapid development and dissemination of hybrids pushing OPVs to marginal areas.

Pearl millet varieties have seed yields two to three times higher than inbred seed parents. The intra-population variability in pearl millet OPVs contributes toward greater resilience to several biotic and abiotic types of stress in contrast to the single-cross hybrids of pearl millet, which, when developed initially,

were highly vulnerable to the downy mildew epidemic. The genetic heterogeneity of OPVs confers durable resistance to downy mildew, and a variation in flowering confers pollen-based escape from ergot and smut infection, which contrast well with frequent downy mildew epidemics and their greater vulnerability to ergot and smut. Consequently, improved pearl millet OPVs are readily acceptable to farmers and are easier to multiply, and hence have carved a niche for themselves even in India, where hybrids are the preferred cultivars (Sanjana Reddy, 2017). The genetic improvement of OPVs started in the 1930s and could not progress beyond a certain limit because of a narrow genetic base. During the 1970s, with the introgression of African germplasm lines especially from Western Africa, the primary center of diversity was led by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) to enhance the genetic diversity of this crop. Due to such focused efforts, composites and OPVs were developed largely based on these germplasm lines. These populations were also a source for the breeding lines, which were widely used over the years by both the public and private sectors. The *Iniadi* germplasm, acquired from western Africa, has been extensively used in India, the USA, and other places worldwide, and several high-yielding hybrids were derived from it. OPV ICTP 8203 is also based on the *iniadi* germplasm. Apart from hybrids, OPVs such as WC-C75, Raj 171, ICMV 155, ICMV221, ICTP 8203, CZP 9802, and JBV 2 became very popular with farmers soon after their release (Yadav and Rai, 2013). In Maharashtra, a substantial proportion is still an OPV (variety ICTP 8203). In the arid tracts of west Rajasthan, landraces/OPVs are widely grown because of an extremely risky production environment. Though research focus has shifted to hybrids, the development of OPVs with good yield potential is possible. Hybrid parents with improved resistance to downy mildew and with good yield levels were further derived and formed the background of modern-day hybrids. These breeding lines were used worldwide, predominantly in India (Rai et al., 2014). However, the wide usage of a single source of germplasm, such as *iniadi*, poses a threat of disrupting existing heterotic patterns,

TABLE 1 | Varieties used in the study.

Variety	Details	Testing zone
Test entries		
MP 590 (VPMV 9)	Test variety bred at AICRP-PM, Vijayapura, Karnataka	A
MP 595 (GBL 2)	Test variety bred at AICRP-PAU, Ludhiana, Punjab	A, B
MP 596 (GBL 5)	Test variety bred at AICRP-PAU, Ludhiana, Punjab	A, B
MP 597 (HBC 53)	Test variety bred at AICRP-Hisar, Haryana	A, B
MP 598 (PC 720)	Test variety bred at ICAR-IARI, New Delhi	A, B
MP 599 (PC 721)	Test variety bred at ICAR-IARI, New Delhi	A, B
MP 600 (PC 722)	Test variety bred at ICAR-IARI, New Delhi	A, B
Checks		
PC 383	Bred at ICAR-IARI, New Delhi, released in 2001 for cultivation in A-zone	A
PC 701	Bred at ICAR-IARI, New Delhi, released in 2016 for cultivation in A-zone	A
JBV 2	Bred at AICPMIP Gwalior and ICRISAT, released in 1999 for cultivation in A-zone	A
ABV 04	Bred at ANGRAU, Ananthapuramu, released in 2019 for cultivation in B-zone	B
ICMV 155	Bred at ICRISAT, released in 1991 for cultivation in B-zone	B
PC 612	Bred at ICAR-IARI, New Delhi, released in 2011 for cultivation in B-zone	B
Raj 171	Bred at AICPMIP ARS Durgapura, Jaipur and ICRISAT, released in 1992 for cultivation across India	A, B
Dhanshakti (ICTP 8203 Fe)	Bred at ICRISAT, released in 2014 for cultivation across India	A, B
ICMV 221	Bred at ICRISAT, released in 1993 for cultivation across India	A, B

which may be noticed in upcoming years. Diversification and development of new OPVs can also be used as a continuous source of variability for the generation of hybrid parental lines. Since OPVs are more preferred by resource-poor farmers in marginal areas, the nutritional status of OPVs with optimum levels of Fe and Zn has to be maintained, as the grain is consumed at the source. In this context, it becomes pertinent that the existing populations are evaluated for yield potential, adaptability, and nutritional status, and draw conclusive steps for the breeding and testing of new OPVs.

The research on Pearl millet improvement in India is carried out by the All India Coordinated Research Project on Pearl Millet (AICRP-PM), administered by the Indian Council of Agricultural Research (ICAR) through a network of 13 AICRP-PM centers and several voluntary ones. Newly developed cultivars are tested in multiple locations to determine their stability and performance before commercial release. Interpreting the genotype-by-environment interaction (GE) is essential for the identification of stable genotypes across environments to thereby obtain a correct ranking of genotypes and identify ideal genotypes for the target environment and ideal environment for discriminating genotypes. Several statistical methods are available for the study of GE. Among them, the additive main effects and multiplicative interaction (AMMI) and GGE biplots are frequently used for multi-environment trial (MET) data analysis. The GGE biplot analysis proposed by Yan et al. (2000) considers both genotype main effects and GEI effects for the analysis (Miranda et al., 2009), while genotype effects are not considered in AMMI. Therefore, the GGE biplot model is considered as an efficient method for identifying the best genotypes and test environments (Ding et al., 2007).

The current study was accomplished with the involvement of 20 locations across India to (i) identify best representative/

discriminating locations for the evaluation of pearl millet OPVs, (ii) identify stable OPVs across locations and those suitable to specific zones, and (iii) determine variation in grain Fe and Zn contents across locations in the OPVs.

MATERIALS AND METHODS

Plant Material

The experimental material consisted of seven test or new varieties and nine released and popular OPVs, including biofortified variety Dhanshakti, used as checks in a population trial. Of them, seven test varieties and six checks were evaluated across 10 locations in zones A₁ and A (pooled as A, referring to northern India), while six test varieties (new OPVs) and six checks were evaluated across 10 locations in zone B. For pooled analysis, six test varieties evaluated across 20 locations in the A and B zones, along with three common checks and additional three zonal-adapted checks in respective zones, during the 2019 rainy season were used. The information that pertains to the varieties used in this study is presented in **Table 1**.

Test Locations and Experiment

The multi-location testing was done at 20 locations in 11 states. The states of Maharashtra, Andhra Pradesh, and Rajasthan had three locations; the states of Karnataka, Haryana, and Madhya Pradesh were represented with two locations; and the states of Tamil Nadu, Telangana, Gujarat, Punjab, and Delhi were represented with a location each. Detailed features of these test locations and dates of sowing are given in **Table 2**. The crops were sown with the onset of monsoon in each of these locations. In each location, the experiment was conducted in a randomized complete block design with three replications. The plot size of each genotype varied from 12 to 14.4 m² across locations,

TABLE 2 | Test locations used in the study.

Location	Soil	pH	Date of sowing	Latitude	Longitude	Altitude
Zone A						
Mandor	Sandy loam	8.1	22 July	26°34' N	73°05' E	241 m
Bikaner	Sandy	8.0	25 July	28°01' N	73°30' E	250 m
Durgapura	Sandy loam	–	8 July	26°90' N	75°80' E	425 m
Jamnagar	Medium black	7.6	2 August	22°28' N	70°05' E	17 m
Hisar	Sandy loam	–	27 June	29°15' N	75°70' E	215 m
Gurugram	–	–	5 July	28°46' N	77°03' E	217 m
Gwalior	Sandy loam	7.5	13 July	26°22' N	78°18' E	211 m
Morena	Clay loam	7.3	22 June	26°49' N	77°99' E	177 m
Ludhiana	Sandy loam	8.0	18 July	30°90' N	75°85' E	262 m
New Delhi	Sandy loam	7.8	10 July	28°61' N	77°20' E	216 m
Zone B						
Aurangabad	Medium black	7.5	2 July	19°86' N	75°30' E	568 m
Niphad	Medium black	8.8	13 July	20°08' N	74°10' E	569 m
Dhule	Medium black	8.6	7 July	21°08' N	74°01' E	210 m
Vijayapura	Shallow black	8.7	21 July	16°50' N	75°43' E	594 m
Malnoor	Medium black	8.2	20 July	17°03' N	76°15' E	460 m
Ananthapuramu	Red sandy loam	6.5	9 August	14°68' N	77°60' E	335 m
Perumallapalle	Red sandy	6.9	15 July	13°39' N	79°35' E	183 m
Vizianagaram	Red sandy loam	6.4	15 July	18°10' N	83°39' E	74 m
Palem	Sandy loam	7.9	27 June	16°35' N	78°10' E	642 m
Coimbatore	Clay loam	7.8	2 July	11°02' N	76°96' E	427 m

which included approximately 100 plants. Plot yield data were converted to tons per hectare using the plot size as a factor.

Trait Measurements

Eight yield-related traits were measured in each trial. The flowering time (DF) was measured as the number of days taken from the date of sowing to the date on which 50% of plants in a plot showed full stigma emergence. The length in centimeters of fully matured plants from the base of the plant to the top of the ear head was recorded as plant height (PHT). The number of productive tillers (NPT) was counted as the total number of tillers that bear ear head with grains per plant. The panicle length (PL) was measured from the base of the panicle to its tip and recorded in centimeters. The panicle diameter (PD) was measured at the maximum thickness of the panicle in centimeters. Grain yield (GY) was estimated by weighing the grains obtained after drying and threshing of the panicles at 12% moisture content and expressed in grams. Then, the weight per plot was extrapolated into t/ha. For measuring dry fodder yield (DFY), harvested plants were allowed to sun-dry for 7–10 days. The weight recorded per plot was extrapolated to t/ha. For measuring 1,000-grain weight (1,000 GWT), a sample of 1,000 grains was counted randomly from the threshed seed, and the weight was recorded in grams.

Micronutrient Analysis

The grain Fe and Zn contents were analyzed using an energy-dispersive x-ray fluorescence spectrometry machine (ED-XRF), model X-Supreme 8000 from OXFORD, installed in the Pearl

Millet Breeding program at ICRISAT, Patancheru, India. The ED-XRF method for pearl millet established and reported a higher correlation between ICP-OES and ED-XRF ($r = 0.9$) for both Fe and Zn, as suggested by Govindaraj et al. (2016). The quantified grain iron and zinc levels were measured in milligrams per kilogram (mg kg^{-1}) of the seed and interpreted in the same unit. All possible care was taken from sampling to laboratory analysis to avoid any contamination.

Statistical Analysis

Recorded data from eight yields and two quality traits were subjected to combined analysis of variance (ANOVA) to investigate genotypes (G), environments (E) and genotype × environment interaction (GEI) effects using GenStat 12th edition. In the combined ANOVA, the genotypes were considered as fixed effects, while the environments and replications were considered as random effects. As the GEI was significant, GGE biplot method (Yan et al., 2000) was employed to analyze GE interaction and assess the stability of GY, DFY, Fe and Zn data, and the pattern of response of OPVs tested in 20 locations. For the eight individual trials, the Pearson's correlation coefficients were calculated using R-software (R Development Core Team, 2019). The repeatability of a variety trial is derived as the proportion of the variation due to genotype effects. Variance components and heritability across the locations were estimated. The broad-sense heritability was calculated as: $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{gl}^2/l + \sigma_{e/lr})$, where σ_g^2 is the genotypic variance, σ_{gl}^2 is the interaction variance of genotype

TABLE 3 | Estimation of variance components and broad sense heritability (h^2_{bs}) for yield, quality, and agronomic traits in A-zone, B-zone, and across all locations in India.

Statistics	GY	DFY	DF	PHT	NPT	PL	PD	1,000 GWT	Fe	Zn
A-zone										
h^2_{bs}	0.79	0.68	0.84	0.9	0.5	0.83	0.82	0.75	0.93	0.83
Genotype variance	2.1	36.7	112.4	6517.2	1.0	118.6	0.9	8.0	2053.3	345.1
Location variance	8.8	1244.3	633.0	43850.9	15.0	389.6	3.6	31.2	2913.9	3331.6
Genotype × location variance	0.4	11.9	17.8	639.0	0.5	19.6	0.2	2.0	134.4	58.1
Residual variance	0.1	1.4	1.9	165.6	0.1	2.7	0.0	0.3	83.8	30.4
LSD ($P < 0.05$)	0.4	1.9	2.2	20.7	0.5	2.7	0.3	0.8	14.7	8.9
CV (%)	11.1	14.5	2.8	6.1	14.2	6.0	7.5	5.5	16.8	14.9
B-zone										
h^2_{bs}	0.74	0.62	0.9	0.91	0.71	0.89	0.91	0.82	0.97	0.86
Genotype variance	1.9	13.8	144.5	5294.2	2.0	92.6	1.6	28.2	2507.4	336.0
Location variance	28.2	120.0	289.9	32110.1	13.5	169.3	1.8	153.7	3126.9	1588.7
Genotype × location variance	0.5	5.3	14.5	476.8	0.6	10.6	0.1	5.0	69.9	46.0
Residual variance	0.1	0.6	2.2	104.0	0.2	3.1	0.0	0.7	19.1	16.1
LSD ($P < 0.05$)	0.5	1.2	2.4	16.4	0.8	2.8	0.3	1.3	7.0	6.5
CV (%)	15.9	14.7	3.1	5.5	18.5	7.2	7.7	6.7	9.4	12.9
Pooled										
h^2_{bs}	0.88	0.79	0.93	0.94	0.73	0.9	0.92	0.89	0.98	0.93
Genotype variance	4.1	41.5	326.5	10873.7	1.8	198.9	2.1	33.7	5651.7	816.1
Location variance	13.0	485.0	682.3	28085.6	9.4	272.4	2.1	116.4	2727.3	1992.9
Genotype × location variance	0.5	8.5	23.1	704.4	0.5	19.3	0.2	3.6	114.8	53.2
Residual variance	0.1	0.8	2.1	146.1	0.1	3.3	0.0	0.5	56.7	24.4
LSD ($P < 0.05$)	0.4	1.5	2.4	19.4	0.6	2.9	0.3	0.4	12.1	7.9
CV (%)	13.6	14.3	3.0	6.2	16.0	7.1	7.2	13.6	14.6	14.2

with location, σ^2_e was the error variance, l was the number of locations, and r was the number of replicates. The estimates of σ^2_g , σ^2_{gl} , σ^2_e were obtained from an ANOVA with the environment considered as a random effect, as mentioned by Xie et al. (2020).

The grain yield stability of OPVs and suitability of test environments is tested using the GGE biplot based on the following model proposed by Santos et al. (2019).

$$Y_{ij} - \bar{Y}_j = \lambda_1 \xi_{i1} \eta_{1j} + \lambda_2 \xi_{i2} \eta_{2j} + \epsilon_{ij}$$

where Y_{ij} is the mean grain yield of genotype i in environment j ; \bar{Y}_j is the mean grain yield of environment j ; λ_1 and λ_2 are the singular values of the first and second principal components, PC1 and PC2, respectively; ξ_{i1} and ξ_{i2} are the scores of genotype i for PC1 and PC2, respectively; η_{1j} and η_{2j} are the scores of environment j for PC1 and PC2, respectively; and ϵ_{ij} is the error associated with the model.

The genotype-centered and the environment-centered singular value partitioning (SVP) are used for the evaluation of genotypes and environments, respectively (Yan et al., 2011), but symmetric scaling is preferred for the study of which-won-where pattern (Yan, 2002). Genotype-by-trait biplot (GT biplot) is generated from combined data using “Genotype-by-trait biplots

Scaling = 1 option” of GGE biplot software. Here, traits were considered as “tester.” “Which is best for what” analysis is performed to identify the genotypes superior for particular traits. GGE Biplot analyses were performed using the R statistical software, version 4.0.0 (R Development Core Team, 2019) and GGEbiplot ver. 8.2 (Yan, 2001).

RESULTS

Analysis of Variance

ANOVA was performed zone-wise as well as pooled over all the locations. The combined ANOVA across environments evidenced highly significant differences among genotypes for all the recorded traits. The proportion of genotype to GE variance was 4–5 times for NPT and DFY; 9–10 times for GY, PL, and 1,000 GWT; 12–15 times for DF, PHT, PW, and Zn; and 49 times for Fe. The proportion of genotypic variance to total variance was marginally higher for traits GY and 1,000 GWT in the A-zone in contrast to the B-zone; while for other traits, the genotypic variance was marginally superior in the B-zone. The broad-sense heritability estimates for NPT (0.73) and DFY (0.79) were lower than those of the other traits (0.88–0.98). The repeatability of the trial, as measured by broad-sense heritability, was marginally higher for traits GY and DFY in the A-zone; while for the other traits, B-zone estimates were marginally higher (Table 3).

TABLE 4 | Performance of OPVs for grain yield and quality traits pooled across locations.

Varieties/checks		GY	DFY	DF	PHT	NPT	PL	PD	1,000 GWT	Fe	Zn
MP 595	A	2.08	9.42	51.0	241.9	2.7	29.7	2.4	8.3	49.4	38.3
	B	1.96	4.38	48.3	201.1	2.7	25.4	2.2	10.8	44.0	34.0
	AB	2.02	6.90	50.4	221.2	2.6	27.5	2.3	9.5	47.1	36.4
MP 596	A	2.31	8.23	50.9	197.7	2.1	26.6	2.9	9.6	51.7	32.3
	B	2.35	5.09	49.5	176.3	2.6	23.2	2.8	12.5	45.0	29.0
	AB	2.33	6.66	50.5	185.0	2.3	25.0	2.7	11.0	48.4	30.7
MP 597	A	1.73	6.89	48.6	196.7	2.5	27.7	2.8	9.3	59.5	38.6
	B	1.97	4.67	50.0	160.1	2.6	25.0	2.8	13.1	46.0	32.0
	AB	1.85	5.78	49.9	182.6	2.5	26.0	2.8	11.1	53.1	35.5
MP 598	A	2.06	8.04	47.7	212.3	2.2	28.0	2.6	9.1	46.8	33.1
	B	2.17	5.60	47.6	186.5	2.4	25.3	2.7	12.4	41.0	28.0
	AB	2.12	6.82	48.2	201.2	2.3	27.0	2.6	10.7	44.2	30.7
MP 599	A	2.18	9.01	49.8	208.8	2.2	26.8	2.8	9.3	49.9	35.8
	B	2.34	5.82	48.7	195.8	2.4	25.5	2.7	12.0	39.0	29.0
	AB	2.26	7.41	49.4	201.8	2.3	25.9	2.8	10.6	45.0	32.5
MP 600	A	2.25	8.74	48.5	215.6	2.3	28.9	2.7	9.2	49.0	35.7
	B	2.22	5.96	48.0	197.0	2.4	24.7	2.6	12.1	41.0	28.0
	AB	2.24	7.35	48.9	207.2	2.4	27.0	2.7	10.6	45.1	31.9
Mean of new OPVs	A	2.10	8.39	49.4	212.2	2.3	28.0	2.7	9.1	51.1	35.6
	B	2.17	5.25	48.7	186.1	2.5	24.9	2.6	12.2	42.7	30.0
	AB	2.14	6.82	49.5	199.8	2.4	26.4	2.7	10.6	47.2	33.0
Raj 171	A	1.69	6.02	49.2	208.0	2.5	28.4	2.4	8.7	51.9	39.7
	B	1.75	4.28	49.4	191.1	3.0	24.7	2.3	10.6	42.0	34.0
	AB	1.72	5.15	49.9	200.9	2.6	26.7	2.4	9.6	47.1	36.8
Dhanshakti	A	1.58	7.35	45.8	186.9	2.0	25.4	2.8	10.2	74.3	44.3
	B	1.69	4.91	42.7	170.7	2.4	20.7	2.9	13.6	70.0	38.0
	AB	1.63	6.13	44.4	180.8	2.2	23.3	2.8	11.8	72.1	41.4
ICMV 221	A	1.69	6.47	44.2	194.8	2.1	22.6	2.8	9.8	67.3	40.8
	B	1.78	4.23	44.6	175.7	2.3	21.8	2.9	12.9	59.0	35.0
	AB	1.74	5.35	44.7	187.5	2.2	22.2	2.8	11.3	63.4	38.0
Mean of checks	A	1.65	6.61	46.4	196.6	2.2	25.5	2.7	9.6	64.5	41.6
	B	1.74	4.47	45.6	179.2	2.6	22.4	2.7	12.4	57.0	35.7
	AB	1.70	5.54	46.3	189.7	2.3	24.1	2.7	10.9	60.9	38.7
Trial mean		1.99	6.39	48.5	196.5	2.4	25.6	2.7	10.7	51.7	34.9
Lsd (5%) b/w entries		0.10	0.33	0.5	4.3	0.1	0.7	0.1	0.2	2.7	1.8

GY, grain yield (t/ha); DFY, dry fodder yield (t/ha); DF, days to 50% flowering; PHT, plant height (cm); NPT, number of productive tillers; PL, panicle length (cm); PD, panicle width (cm); 1,000 GWT, 1,000 grain weight (g); Fe, Fe (mg/kg); Zn, Zn (mg/kg). Mean values are in bold font.

Mean Performance

The OPVs had better performance for DFY, PHT, PL, Fe, and Zn in the A-zone; GY, NPT, and 1,000 GWT in the B-zone, while no difference was observed for DF and PD across zones. The new OPVs had 27% higher grain yield in the A zone and 25% higher in the B zone compared with the checks. In both zones, varieties MP 596, MP 599, and MP 600 had the highest grain yield (2.18–2.35 t/ha). Similarly, for DFY, new varieties were superior to checks by 27% in the A-zone and 23% the in B-zone. OPVs MP 595, MP 599, and MP 600 had a higher DFY of 8.74–9.42 t/ha in the A zone; while OPVs MP 598, MP 599, and MP 600 had a higher DFY of 5.6–5.96 t/ha in the B zone.

Thus, the two OPVs, MP 599 and MP 600, were superior for both GY and DFY. The biofortified variety, Dhanshakti, had the highest levels of Fe (72.1 ppm) and Zn (41.4 ppm), followed by check ICMV 221. The new OPVs had lesser levels of Fe (21% in A-zone, 25% in B-zone), and Zn (14% in A-zone, 16% in B-zone) compared with checks. Among the new OPVs, MP 597 had higher levels of Fe (53.1 ppm) and Zn (35.5 ppm). When the other agronomic traits were observed across the zones, the new OPVs were late by 3 days, had a taller height of 7–15 cm, longer panicle length of 2.5 cm, and smaller grain size by 0.2–0.5 g/1,000 grains. Minor differences were observed for PD and NPT (Table 4).

TABLE 5 | Location and zonal means for grain yield and quality traits pooled across OPVs.

Location	GY	DFY	DF	PHT	NPT	PL	PD	1,000 GWT	Fe	Zn
Bikaner	1.69	6.78	46.8	168.9	2.0	24.1	2.6	8.9	61.9	42.6
Durgapura	1.93	2.49	48.7	171.5	2.1	25.8	2.3	9.1	51.8	34.7
Gurugram	2.47	17.36	49.9	248.2	1.0	34.8	3.0	8.3	53.9	38.9
Gwalior	2.01	7.86	47.9	224.5	2.4	25.5	2.3	9.2	53.5	19.9
Hisar	2.30	6.33	53.4	257.4	2.9	29.0	2.9	9.3	61.3	48.3
Jamnagar	0.96	2.75	43.6	177.6	1.8	25.0	2.6	8.1	52.7	33.4
Ludhiana	2.62	12.07	51.9	212.7	2.7	29.5	3.2	10.0	40.4	33.2
Mandor	1.60	2.08	48.7	171.5	2.1	25.8	2.3	9.1	50.3	34.6
Morena	1.81	4.72	40.6	224.1	3.3	26.4	2.8	9.5	–	–
New Delhi	2.13	15.52	52.7	213.3	2.6	25.4	2.8	11.5	74.1	53.1
A-zone mean	1.95	7.80	48.42	207.0	2.3	27.1	2.7	9.3	55.2	37.4
Ananthapuramu	2.82	3.85	45.4	178.7	2.9	25.2	2.5	14.9	47.2	25.8
Aurangabad	3.15	4.66	51.3	173.7	1.9	23.0	2.7	13.0	52.0	37.0
Coimbatore	2.31	4.77	46.0	213.2	3.3	23.9	2.5	11.0	31.1	40.0
Dhule	2.23	5.00	51.3	224.2	2.5	26.4	3.1	10.2	64.5	39.2
Malnoor	1.81	5.89	48.3	179.3	3.3	24.3	2.7	11.7	51.4	28.4
Niphad	0.13	3.61	63.3	165.7	2.3	21.0	2.3	10.7	–	–
Palem	1.12	2.15	43.7	161.6	2.6	20.2	2.3	10.6	43.0	22.7
Perumallapalle	1.85	9.29	48.3	204.3	1.6	26.7	2.7	11.8	42.9	29.4
Vijayapura	2.10	6.55	46.7	135.2	2.0	22.6	2.8	11.3	47.3	31.7
Vizianagaram	2.74	4.15	40.9	224.0	2.2	27.7	2.9	15.6	–	–
B-zone mean	2.03	4.99	48.5	186.0	2.5	24.1	2.7	12.1	48.3	32.4
Isd ($P < 0.05$) across locations	0.14	0.49	0.8	6.5	0.2	1.0	0.1	0.4	4.0	2.6

GY, grain yield (t/ha); DFY, dry fodder yield (t/ha); DF, days to 50% flowering; PHT, plant height (cm); NPT, number of productive tillers; PL, panicle length (cm); PD, panicle width (cm); 1,000 GWT, 1,000 grain weight (g); Fe, Fe (mg/kg); Zn, Zn (mg/kg).

Among the A-zone locations, Ludhiana, Gurugram, and Hisar were more productive for GY (2.3–2.6 t ha⁻¹) while Gurugram, Ludhiana, and New Delhi were productive for DFY (12.1–17.4 t ha⁻¹). The grains harvested from New Delhi, Hisar, and Bikaner had higher levels of Fe and Zn contents. In the B-zone, the Aurangabad, Ananthapuramu, and Vizianagaram locations had higher grain yields (2.7–3.2 t ha⁻¹), while the Perumallapalle, Vijayapura, and Malnoor locations had higher fodder yields (5.9–9.3 t ha⁻¹). The grains harvested from the Dhule, Aurangabad, and Malnoor locations had higher levels of Fe and Zn contents (Table 5).

Trait Associations

Increase in grain yield was significantly associated with the enhancement of DFY. However, the grain Fe and Zn contents decreased with an increase in grain yield. Grain Fe content was more in early flowering OPVs with shorter panicles and bigger seed sizes. High Fe content is significantly related to high Zn content in the grain (Figure 1).

Mean Performance and Stability Visualized Through Genotype Main Effect Plus Genotype by Interaction Biplot

The environment-centered (centering = 2) genotype-metric (SVP = 1) biplots without scaling (scaling = 0) for grain yield and fodder yield, economically important traits and Fe and Zn, and grain quality traits are presented in Figures 2A–D, respectively. The first two PCs explained the 67.8% variation for GY, 85.5% for DFY, 90.4% for Fe, and 78.7% for Zn. The AEC abscissa passes through the biplot origin and acts as a marker for the average environment and points toward higher mean values (Yan, 2001). The perpendicular lines to the AEC passing through the biplot origin are referred to as AEC ordinate. These ordinates are depicted as dotted lines in Figures 2A–D. The greater the absolute length of the projection of a cultivar, the less stable it is. Furthermore, the average yield of genotypes is approximated by the projections of their markers to the AEC abscissa (Kaya et al., 2006). Accordingly, MP 596 was the best performing genotype in terms of grain yield, followed by MP 599 and MP 600; while ICMV 221, Raj 171, and Dhanshakti were limited by lower yields. They were also least stable for grain yield with

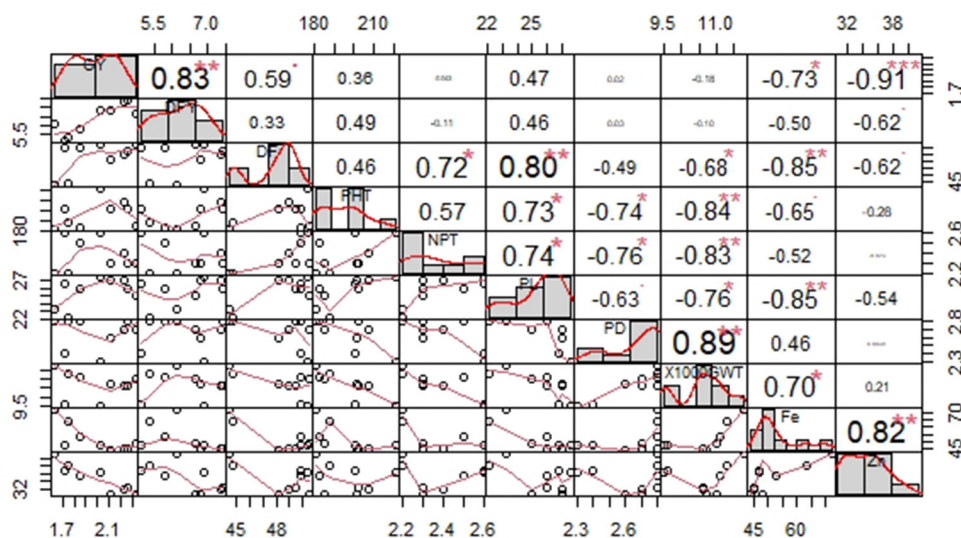


FIGURE 1 | Correlation among yield, quality, and agronomic traits recorded on nine OPVs over 20 locations. Significant effects at * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

higher projection from the AEC abscissa. OPV MP 600 was the most stable among the high-yielding OPVs (**Figure 2A**). For dry fodder yield, MP 599 and MP 600 had the highest yields, but MP 599 was more stable. Compared to the test varieties, the check varieties had a low fodder yield (**Figure 2B**). Dhanshakti had higher Fe (72.1 mg/ka) and Zn (41.4 mgkg⁻¹) contents, followed by ICMV 221 (63.4 mgkg⁻¹ Fe and 38 mgkg⁻¹ Zn). However, ICMV 221 had greater stability for the traits. The rest of the OPVs did not perform well for grain quality traits (**Figures 2C,D**).

Relationship Among Environments

The relationships among the test environments were studied by environment centered (centering = 2), environment metric (SPV = 2), and without scaling (scaling = 0). Combined ANOVA for grain yield (**Figure 3A**) showed that the majority of the angles between their vectors are acute. Acute vector angles are indicative of a closer relationship among the environments (Yan and Tinker, 2006). Thus, the majority of the locations were highly correlated except for the Vizianagaram and Perumallapalle locations and for the Ludhiana, Vijayapura, Mandor, and Durgapura locations, which shows no relationship among them as the angle was 90°. The distance between two environments measures their ability to discriminate genotypes. Thus, the 20 locations could be divided into three groups for grain yield; one with Durgapura, Mandor, Vijayapura, and Ludhiana; second with Vizianagaram and Perumallapalle; and the other 14 locations forming the third group. The groupings did not correlate with the A-zone and B-zone groupings that exist or with geographical identity. The environments were diverse with respect to fodder yield. The locations Perumallapalle, Niphad, Ludhiana, Malnoor, Gwalior, Vijayapura, New Delhi, Dhule, Mandor, and Durgapura were

related for DFY (**Figure 3B**). All of the locations were highly correlated for grain Fe content (**Figure 3C**), while the majority of the locations were correlated for grain Zn content (**Figure 3D**) with the exception between Gwalior and Hisar, which has a right angle between them showing no-relationship.

In **Figure 4**, “average environment” is represented by a small circle on the average environment axes (AEA). The length of environmental vectors is proportional to the standard deviation of the genotypes in the environments. The longer environmental vectors indicate that the environment is more differentiating for the trait among the genotypes. Another important criterion in evaluating environments is the test of their representativeness. The average environment coordination (AEC) line crosses the center of the biplot and the medium environment, and the angle of each vector with the AEC axis is a criterion for identifying the sample environment. Environments with smaller angles with the AEA are most representative of the average test environments. A suitable environment should have two criteria at the same time: distinctive and a target environment. The Ananthapuramu location was closest to the average environment, and thus is the most representative or discriminating environment, followed by the New Delhi location. While ranking the genotypes in near-average environment Ananthapuramu, MP 596, MP 599, and MP 600 had higher GY; MP 595 and MP 598 had moderate yield, and genotypes MP 597, Raj171, ICMV 221, and Dhanshakti had lower than average yield. Variety MP 600 was highly stable, followed by MP 599 (**Figure 5**).

Which Won Where and Mega Environment Identification

A mixture of crossover and non-crossover types of GEI in MET data is of very common occurrence (Kaya et al., 2006; Fan et al.,

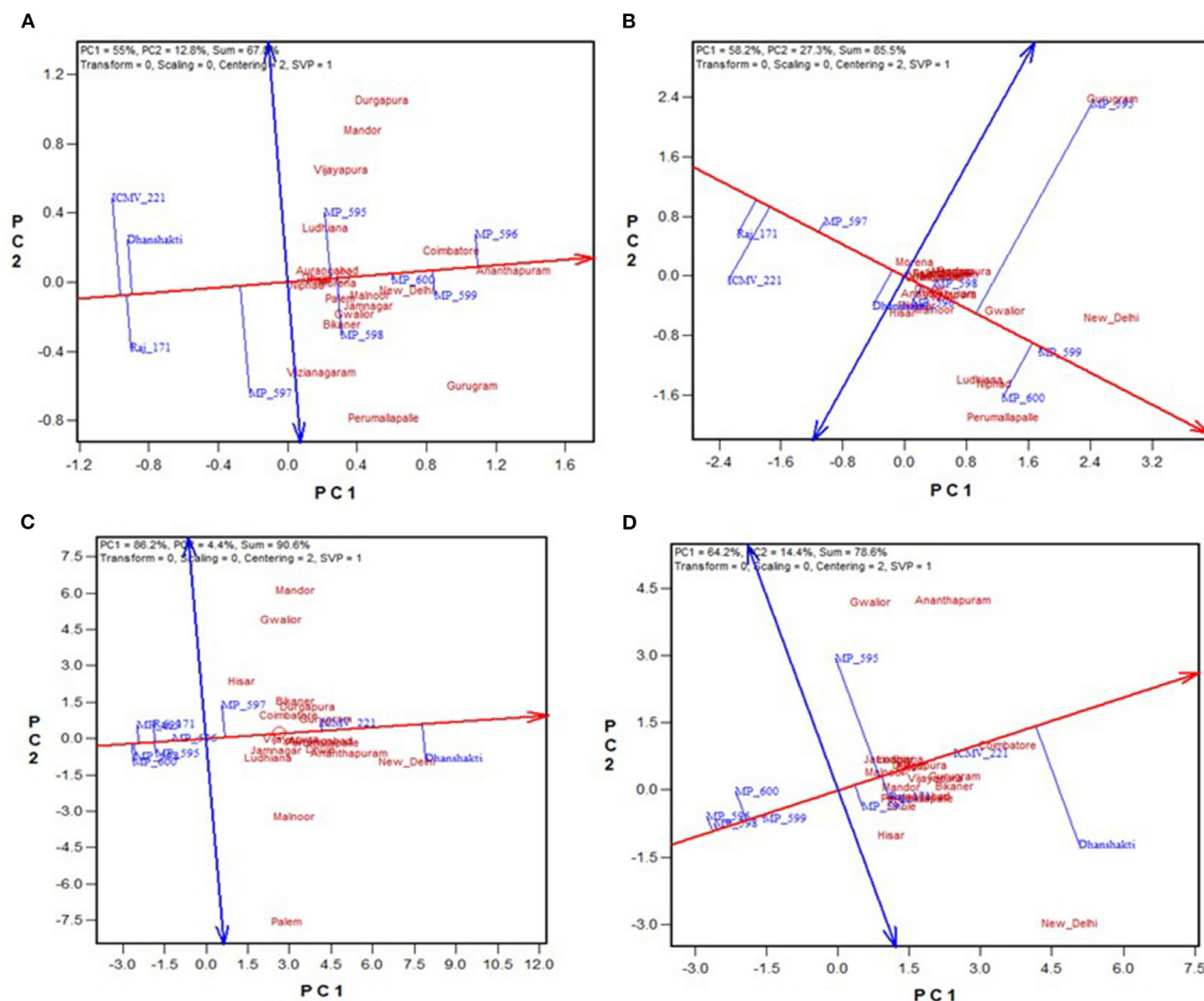


FIGURE 2 | GGE biplot showing “mean vs. stability” of nine pearl millet OPVs across 20 locations for (A) grain yield, (B) dry fodder yield, (C) grain Fe content, and (D) grain Zn content.

2007; Sabaghnia et al., 2008; Rao et al., 2011). The “which-won-where feature” of the GGE biplot graphically addresses crossover GE, mega-environment differentiation, specific adaptation, etc. (Gauch and Zobel, 1997; Yan et al., 2000; Yan and Tinker, 2006; Putto et al., 2008; Rao et al., 2011). The “which-won-where” graph is constructed by joining the farthest genotypes in a polygon. From the origin of the biplot, perpendicular lines, referred to as equity lines, are drawn to the sides of the polygon, separating the polygon into several sectors (Yan, 2001). Genotype at the vertex is the best performing genotype in the environment falling in that sector (Yan and Tinker, 2006). The “which-won-where” biplots for GY and DFY over pooled locations are presented in **Figures 6A,B**. The biplots indicated the existence of crossover GEI and the existence of mega-environments (ME). For grain yield over

pooled locations, the hexagon has six genotypes, MP 595, MP 596, MP 599, MP 597, Raj 171, and ICMV 221 at its vertices. The equity lines divided the biplot into six sectors, of which three retained 20 locations. The testing locations partitioned into three MEs, ME1 with locations Vijayapura, Ludhiana, Dhule, Aurangabad, Morena, Palem, Jamnagar, Malnoor, Hisar, Jamnagar, New Delhi, Ananthapuramu, Mandor, Durgapura, and Coimbatore with MP 595, MP 596, and MP 599 as the winning genotypes. ME2 consisted of locations Vijayanagaram, Perumallapalle, and Nipahad with MP 597 as the winning genotype. ME3 consisted of the Bikaner and Gurugram locations with no genotype performing better for these locations (**Figure 6A**). The correlation among the locations did not exist in terms of geography. For fodder yield, five genotypes were placed at the vertices of the

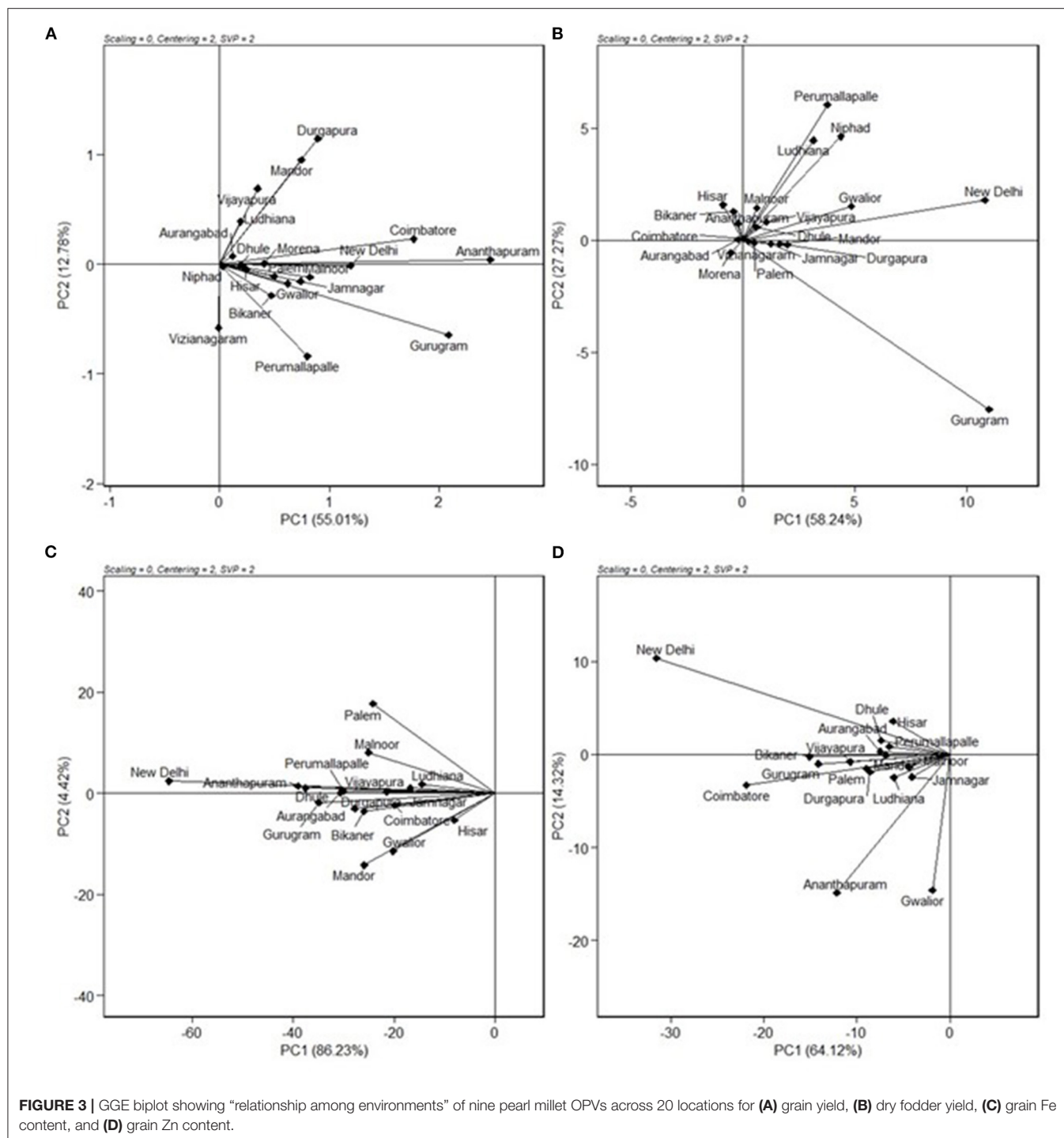


FIGURE 3 | GGE biplot showing “relationship among environments” of nine pearl millet OPVs across 20 locations for **(A)** grain yield, **(B)** dry fodder yield, **(C)** grain Fe content, and **(D)** grain Zn content.

pentagon, and the biplot was divided into five sectors. ME1 had seven locations: Perumallapalle, Niphad, Ludhiana, Malnoor, Vijayapura, Hisar, Bikaner, and Ananthapuramu with MP 599 and MP 600 as winning genotypes. The second largest ME had six locations: Mandor, Durgapura, Jamnagar, Palem, Gurugram, and Dhule with MP 595 as the winning genotype. ME3 had Coimbatore, Aurangabad, Morena, and Vijayanagaram

with Raj171 as the winning genotype; while Gwalior and New Delhi fell into ME4, which had no winning genotype (**Figure 6B**).

“Which is best for what” analysis of the genotype × trait biplot helped to compare genotypes on the basis of multiple traits, and to identify genotypes superior for a particular trait (**Figure 7**). The biplot indicates that MP 595 was the best for

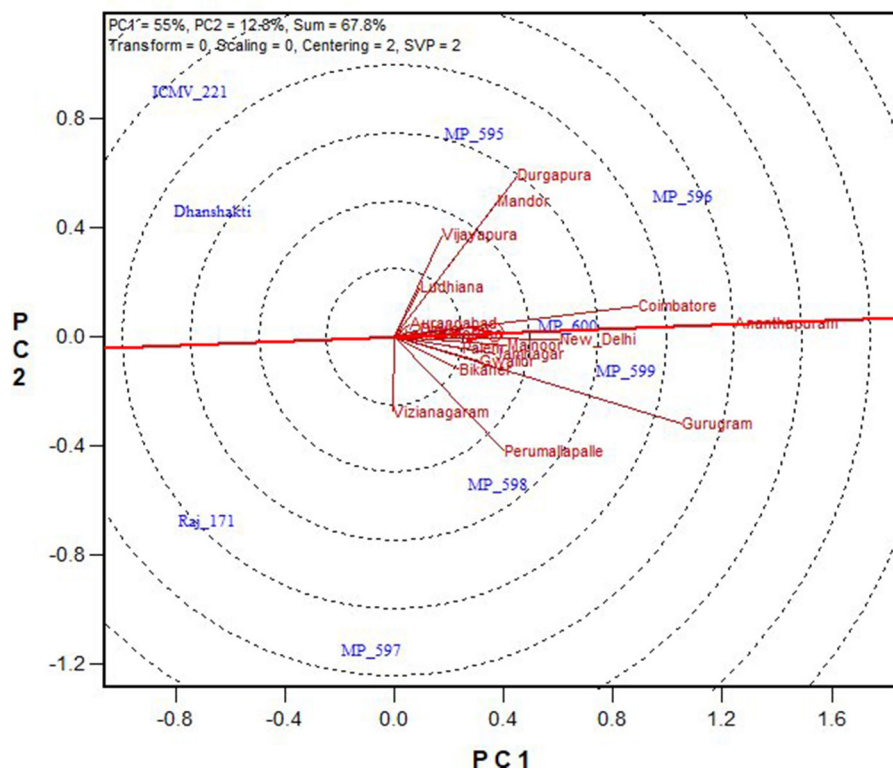


FIGURE 4 | Ranking of environments for grain yield based on discriminating ability and representativeness.

PHT, PL, NPT, and late flowering. OPVs MP 596, MP 598, MP 599, and Mp 600 were better for GY and DFY; while Dhanshakti and ICMV 221 performed well for early flowering, grain Fe and Zn contents.

DISCUSSION

Pearl millet is grown in arid and highly arid tracts of India under minimal or no inputs with hybrids and OPVs as cultivar options. Over a period of time, 61 OPVs are released. This figure is very low compared to the hybrids available for the commercialization of pearl millet. However, few OPVs, such as ICTP 8203 and ICMV 221, are popularly grown, owing to their resilience to marginal conditions and good grain quality apart from reasonable yield levels. Though the OPVs have not received the attention they deserve, few research centers are continuing to develop OPVs with improved yield. Determination of the performance of improved OPVs to compare with that of the popular OPVs for yield and quality traits, and identification of suitable OPVs for target locations are required for focusing research efforts. The 20 multi-location testing sites across India used for the study on pearl millet are handled by AICRP and represent diverse pearl millet production ecosystems. GGE biplot, effectively used in many crops, has been used to analyze

MET data and interpret complex GEI (Yan, 2001; Yan and Tinker, 2006), and to obtain high yielding and stable cultivars, derive the relationship among the environments, identify an ideal environment besides “which won-where,” and delineate mega-environments among the testing locations (Yan et al., 2007). We have studied the GEI among nine pearl millet OPVs (six new and three popular varieties) across 20 locations performing GGE biplot analysis. In the combined data from the current study, environment or location contributed 48–91% of the variation in the data, while the contribution of genotype is from 8 to 66% for 10 traits. The interaction of genotype with location is less (1–4%) though significant (Table 3). Gauch and Zobel (1997) reported that normally in MET data, environment accounts for about 80% of the total variation. In barley MET data, environment accounted for as high as 76.7% (Jalata, 2011). Similar trends are observed in proso millet with environment contributing up to 85% (Pan-pan et al., 2016) and up to 82% in sunflower (Santos et al., 2019). However, Tefera (2018) reported a moderate 51.6% of variation being explained by environment in soybean MET data, and Krishnamurthy et al. (2017) reported 40.5% in rice MET data. In this study, GL explained a lesser proportion of the variation than G alone. This explains the lesser fluctuations among the cultivars with change in environment as they are known for high adaptability and stability. Though low, the significant GE indicates that the genotypes showed varied

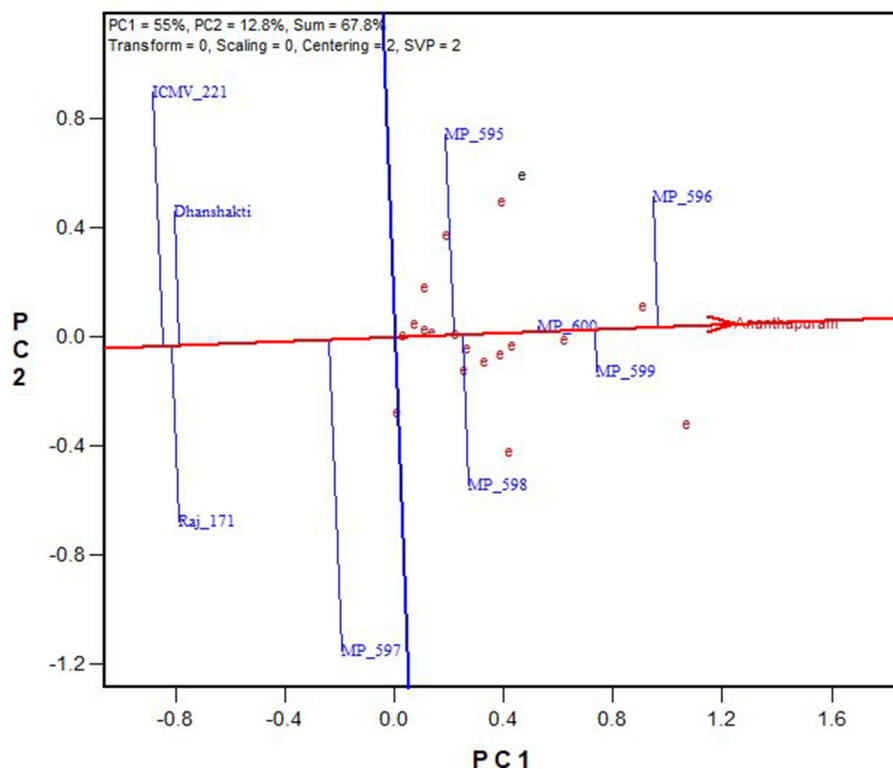


FIGURE 5 | Ranking of genotypes based on their performance for grain yield in near-ideal location, Ananthapuramu.

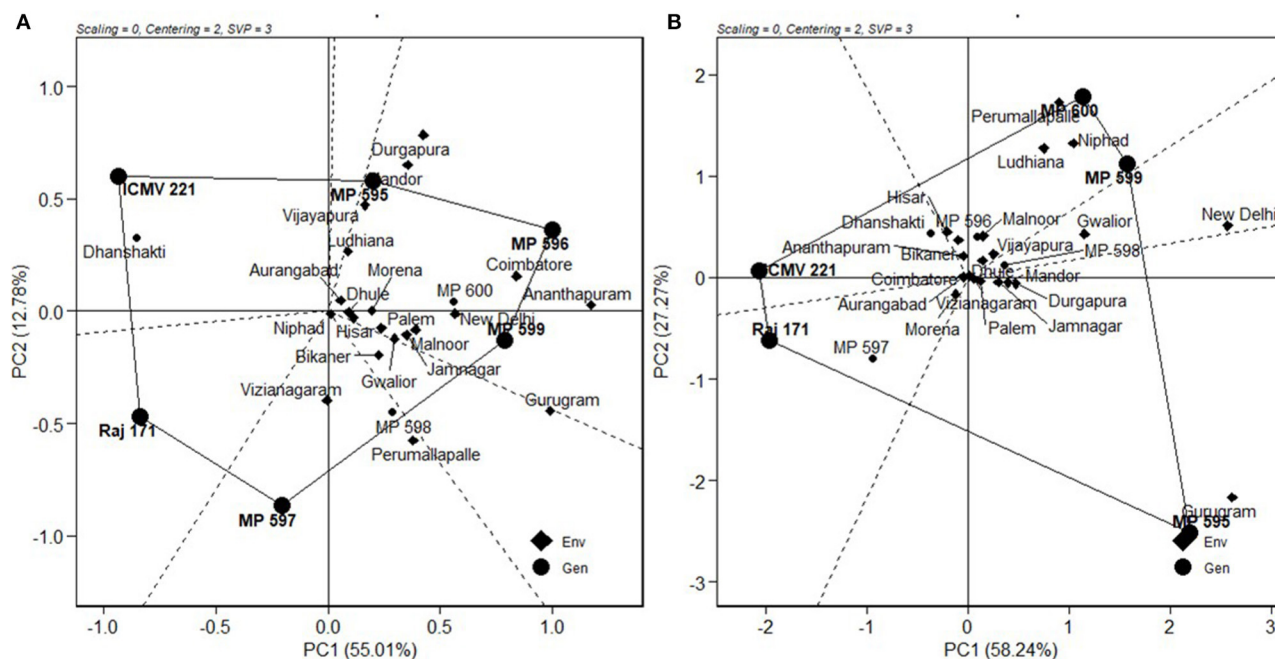


FIGURE 6 | Which-won-where analysis of the genotypes for (A) grain yield and (B) fodder yield.

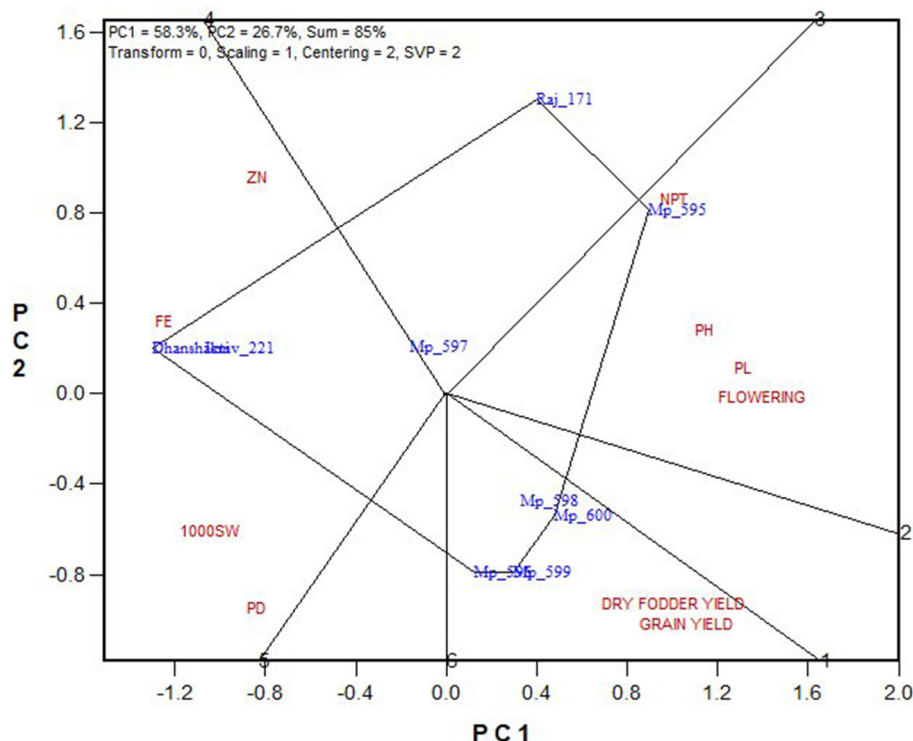


FIGURE 7 | Polygon view of GT biplot indicating which is the best genotype for the target traits.

performances across locations. In the GGE biplot analysis, the complex GEI are simplified in different PCs; and if the first two PCs explain more than 60% of the (G and GL) variability in the data and the combined (G and GL) effect accounts for more than 10% of the total variability, then the biplot adequately approximates the variability in $G \times E$ data (Rakshit et al., 2012). In this study, the first two PCs explained 67.8% variation for grain yield and 85.5% for fodder yield. In addition, **Table 3** indicates that G and GL together accounted for 8 and 14% of total variability for grain and fodder yields, respectively. Thus, the graphical representation of the biplots can be used for deriving stable and ideal genotypes and ideal environments.

Ideal Genotypes

A genotype is considered ideal if it has a high mean yield and less variable across locations and seasons. The quality of the data could be considered quite reliable because of moderate to high broad-sense heritability (73–98%) over locations (**Table 3**). It is evident from **Figure 3A** that the highest grain yielders, MP596 and MP 599, are moderately stable, while MP 600 is comparatively a good yielder with high stability. For fodder yield, MP 599 and MP 600 are good yielders with MP 599 having high stability (**Figure 3B**). The new OPVs performed exceptionally well for mean performance and stability compared with proven and popularly cultivated OPVs, ICTP 8203, ICMV 221, and Raj171 (**Figures 3A,B**). It is a known fact that the OPVs in pearl millet would have more stable yields, are more

widely adapted than hybrids, and are less vulnerable to pests and diseases (Charyulu et al., 2014). Improving the OPVs for grain yield indirectly influences the income sustainability of farmers. Variety MP 600, though highly stable for GY, showed less stability for DFY. Similarly, MP 599 was highly stable for DFY and comparatively less stable for GY. A genotype showing stability for a trait may not necessarily be stable for other traits. As different traits are governed by a different set of genes and the environment influences the overall cumulative expression of different sets of genes, the genotypes vary with yield and stability. Similar observations have been reported by Rakshit et al. (2012) with sorghum. However, other OPVs have followed similar trends in yield and stability for grain and fodder yields. As grain yield and fodder yields are more preferred traits in OPVs by farmers, OPVs MP 599 and MP 600 can be recommended for cultivation across all regions and are identified as ideal genotypes for recommendation to farmers.

Ideal Environments

The “ideal” test environment is that which is most discriminating (brings out the differences among the genotypes) and most representative (represents the target region). Discrimination ability and representativeness of a location can be viewed conveniently from the biplot. The environment with a longer vector and the smallest angle with an ideal environment are identified as a perfect test location in terms of being more discriminating and most representative of overall locations

(Pan-pan et al., 2016). Locations Ananthapuramu, Gurugram, Durgapura, Mandor, and Coimbatore, with more vector lengths, are more discriminating. The near average locations, namely, Ananthapuramu and New Delhi, are more representative and are suitable for selecting more adapted genotypes. On the other hand, Durgapura and Gurugram, being discriminating and non-representative, are useful for selecting specifically adapted genotypes. With the advantage of such graphical representation, where a generally adapted environment and a specific environment can be identified, cultivar choices and breeding schemes can be made. Similar views are put forth by Jalata (2011) and Rakshit et al. (2012). Closer relationships between the test environments indicated that the same information could be obtained from fewer environments. Thus, for initial testing, similar environments may be removed in future multi-location testing of pearl millet cultivars. This also ensures the optimal allocation of scarce resources while formulating MLTs. The absence of wide obtuse angles between environment vectors (**Figure 4**) indicates that there were no negative correlations among the test environments, suggesting the absence of a strong crossover GEI across locations for grain and fodder yields, as suggested by Yan and Tinker (2006). This indicated that genotypes performing better in an environment would also be performing in the same direction in another environment, which means that ranking of genotype does not change from location to location. Even though a mixture of crossover and non-crossover types of GEI in MET data is of very common occurrence (Rao et al., 2011), the data from this study did not show a crossover type of interaction, as the genotypes included in this study are OPVs with inherent resilience, and the check varieties are proven for stability and popularly grown by farmers. Stability is also a response to the environment due to the combined properties of their gene combinations. Being more discriminative and representative among all the testing locations, Ananthapuramu is the ideal environment. Also, in this ideal testing environment, OPVs MP 596, MP 599, and MP 600 have high grain yield.

“Which-won-where” is the most attractive feature of the GGE biplot, which graphically addresses crossover GEI, mega-environment differentiation, specific adaptation, etc., and is widely used by several researchers on many crops (Gauch and Zobel, 1997; Yan and Tinker, 2006; Rao et al., 2011; Krishnamurthy et al., 2017; Jadhav et al., 2019; Pan-pan et al., 2016). Based on this graphical representation, for grain yield, the testing locations were partitioned into three mega environments (ME). ME1 was represented by Vijayapura, Ludhiana, Dhule, Aurangabad, Morena, Palem, Jamnagar, Malnoor, Hisar, Jamnagar, New Delhi, Ananthapuramu, Mandor, Durgapura, and Coimbatore, with MP 595, MP 596, and MP 599 as winning genotypes. For ME1, Ananthapuramu can be selected as the most representative environment. ME2 consisted of locations Vijayanagaram, Perumallapalle, and Nipahad, with MP 597 as the winning genotype and Perumallapalle as the most representative environment. ME3 consisted of the Bikaner and Gurugram locations, for which no genotype performed better. Gurugram can be selected from ME3, as it is more discriminative than Bikaner (**Figure 4**), for initial testing of

cultivars and planning of breeding activities. However, this mega-environment pattern needs to be verified through multi-year and multi-environment trials (Rakshit et al., 2012), as proposed for wheat (Yan et al., 2000). The advantage for grain and fodder yields in the new OPVs compared to traditionally grown check varieties can be clearly observed from the study. Hence, more efforts can be targeted for new OPV development, as the advantage of OPVs in resource-constrained and environment-challenged areas cannot be ignored.

Grain Quality

Genetic variance was very high for grain Fe content and moderate for Zn content, indicating less influence of interaction with the environment on the expression of these traits. The checks had higher grain Fe and Zn content than the new OPVs, while the new OPVs had higher grain and fodder yields. This can also be seen from the trait associations where there was a significantly negative correlation between grain yield and quality traits. However, the new OPVs meet the minimal requirement for grain Fe content of 42 ppm and for grain Zn content of 32 ppm, which is fixed by the AICRP on pearl millet. Though the Fe-rich variety, Dhanshakti, has higher levels of Fe and Zn contents, the check variety, ICMV 221, has comparative levels of Fe and Zn and more stable for grain quality. The grains harvested from the A-zone locations (more from New Delhi, Hisar, and Bikaner) have higher Fe and Zn contents compared with those from the B-zone locations (Dhule, Aurangabad, and Malnoor). The grain Fe and Zn contents seem to be under genetic control and show very less interaction with the environment. Similar to other traits, the influence of location is very high, and the northern part of India has higher levels than the southern India locations. Thus, while fixing the minimal levels of Fe and Zn contents in the grain for varietal release, this aspect has to be considered.

Genotype × Trait Associations

From the trait relationships and GT biplot, it can be observed that the grain quality traits (grain Fe and Zn contents) are highly related and associated with seed size and panicle density along with early flowering and with check varieties ICMV, 221, and Dhanshakti being promising for them. Most of the tested varieties performed best for the economically important traits, GY and DFY, which are highly correlated with each other, poorly correlated with agronomic traits, and negatively correlated with grain quality traits. For other agronomic traits, MP 595 has performed well. Though the grain and dry fodder yields are the most important traits in any crop improvement program, the pearl millet cultivars grown in arid regions should have early flowering and high tillering to sustain harsh climate, apart from yield and quality. The cultivars grown in semi-arid regions require high grain yield contributed through increased panicle length and density in medium to late maturity background. Hence, selection based on correlated response is required to develop a variety that has higher yields in the desired flowering and tillering background along with good grain quality that is not found in a single variety in this study. Thus, this study has identified a wider scope to develop

promising OPVs of pearl millet with breeding strategies based on its outputs.

CONCLUSION

Analysis of variance showed significant differences among genotypes and locations. The GEI was significant though comparatively less than location and genotype effects as the cultivars used in the study are known for their stability. The study has identified the best varieties suitable for cultivation across 20 locations of pearl millet growing areas in India. OPVs MP 599 and MP 600 are identified as ideal genotypes, because they showed higher grain and fodder yields and stability than other cultivars. The study has also shown that the genotypes with high mean performance and stability for a trait may not show similar performance for another trait. On the other hand, the Ananthapuramu location had the best discrimination and better representativeness than other locations. Therefore, Ananthapuramu is the ideal test site for selecting pearl millet cultivars effectively for adaptation across India while Ananthapuramu, Perumallapalle, and Gurugram can be used as initial testing locations based on this study. $G \times E$ interaction had very little influence on grain Fe and Zn contents. Dhanshakti had higher levels of grain Fe and Zn contents, while ICMV 221 was most stable for traits. The check varieties had higher Fe and Zn contents compared with the new OPVs, which

is also reflected in the negative correlation of grain yield with grain Fe and Zn contents. Breeding efforts should be directed to break this linkage. However, the new OPVs meet the required levels of 42 ppm Fe and 32 ppm Zn, fixed by AICRP on pearl millet. Breeders can focus on one of these traits as the correlation between grain Fe and Zn content is very high.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CS and VK conceived and designed the study. HP, PG, LS, KM, SS, RN, HB, KI, MT, DY, RB, AT, VKT, UK, KS, MS, BA, NA, and MG planned and performed the experiments. PS wrote the manuscript. MG, TN, and VAT revised the manuscript. All the authors have read and approved the final manuscript.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the funding of this research from the Indian Council of Agricultural Research—Indian Institute of Millets Research and All India Coordinated Research Project on Pearl Millet.

REFERENCES

- Charyulu, K. D., Bantilan, C., Rajalaxmi, A., Rai, K. N., Yadav, O. P., Gupta, S. K., et al. (2014). *Development and Diffusion of Pearl Millet Improved Cultivars in India: Impact on Growth and Yield Stability*. Working Paper Series No. 52. Patancheru 502 324. Telangana: International Crops Research Institute for the Semi-Arid Tropics, 76.
- Ding, M., Tier, B., and Yan, W. K. (2007). "Application of GGE biplot analysis to evaluate genotype (G), environment (E) and $G \times E$ interaction on *P. radiata*: Case study," in *Australasian Forest Genetics Conference, 11–14, April 2007* (Hobart, TAS: The Old Woolstore).
- Directorate of Millets Development (2020). Available online at: <http://millets.dacfw.nic.in>
- Fan, X. M., Kang, M. S., Chen, H., Zhang, Y., Tan, J., and Xu, C. (2007). Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. *Agron. J.* 99, 220–228. doi: 10.2134/agronj2006.0144
- Gauch, H., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.* 37, 311–326. doi: 10.2135/cropsci1997.0011183X003700020002x
- Govindaraj, M., Rai, K. N., Pfeiffer, W. H., Kanatti, A., and Shivade, H. (2016). Energy-dispersive x-ray fluorescence spectrometry for cost-effective and rapid screening of pearl millet germplasm and breeding lines for grain iron and zinc density. *Commun. Soil Sci. Plant Anal.* 47, 2126–2134. doi: 10.1080/00103624.2016.1228938
- Jadhav, S., Balakrishnan, D., Shankar, G., Beerelli, K., Chandu, G., and Neelamraju, S. (2019). Genotype by environment ($G \times E$) interaction study on yield traits in different maturity groups of Rice. *J. Crop Sci. Biotech.* 22, 425–449. doi: 10.1007/s12892-018-0082-0
- Jalata, Z. (2011). GGE-biplot analysis of multi-environment yield trials of Barley (*Hordeum vulgare* L.) genotypes in Southeastern Ethiopia highlands. *Int. J. Plant Breed. Genet.* 5, 59–75. doi: 10.3923/ijpb.2011.59.75
- Kaya, Y. M., Akcurra, M., and Taner, S. (2006). GGE-biplot analysis of multi-environment yield trials in bread wheat. *Turk. J. Agric. For.* 30, 325–337.
- Krishnamurthy, S. L., Sharma, P. C., Sharma, D. K., Ravikiran, K. T., Singh, Y. P., Mishra, V. K., et al. (2017). Identification of mega-environments and rice genotypes for general and specific adaptation to saline and alkaline stresses in India. *Sci. Rep.* 7:7968. doi: 10.1038/s41598-017-08532-7
- Miranda, G. V., Souza, L. V., Guimarães, L. J. M., Namorato, H., Oliveira, L. R., and Soares, M. O. (2009). Multivariate analyses of genotype \times environment interaction of popcorn. *Pesq. Agropec. Bras.* 44, 45–50. doi: 10.1590/S0100-204X2009000100007
- Pan-pan, Z., Hui, S., Xi-wang, K. E., Xi-jun, J., Li-hua, Y., Yang, L., et al. (2016). GGE biplot analysis of yield stability and test location representativeness in proso millet (*Panicum miliaceum* L.) genotypes. *J. Integr. Agric.* 15, 1218–1227. doi: 10.1016/S2095-3119(15)61157-1
- Putto, W., Patanothai, A., Jogloy, S., and Hoogenboom, G. (2008). Determination of mega-environments for peanut breeding using the CSM-CROPGRO-Peanut model. *Crop Sci.* 48, 973–982. doi: 10.2135/cropsci2007.10.0552
- R Development Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rai, K. N., Gupta, S. K., Govindaraj, M., and Yadav, H. P. (2015). Pearl millet improvement for enhanced productivity – strategies and impact. *Indian Farm.* 65, 19–22.
- Rai, K. N., Gupta, S. K., Sharma, R., Govindaraj, M., Rao, A. S., Shivade, H., et al. (2014). Pearl millet breeding lines developed at ICRISAT: a reservoir of variability and useful source of non-target traits. *SAT e-J.* 1, 1–13.
- Rakshit, S., Ganapathy, K. N., Gomashe, S. S., Rathore, A., Ghorade, R. B., Nagesh Kumar, M. V., et al. (2012). GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica* 185, 465–479. doi: 10.1007/s10681-012-0648-6

- Rao, P. S., Reddy, P. S., Ratore, A., Reddy, B. V. S., and Panwar, S. (2011). Application GGE biplot and AMMI model to evaluate sweet sorghum (*Sorghum bicolor*) hybrids for genotype × environment interaction and seasonal adaptation. *Indian J. Agric. Sci.* 81, 438–444.
- Sabaghnia, N., Dehghani, H., and Sabaghpour, S. H. (2008). Graphic analysis of genotype by environment interaction for lentil yield in Iran. *Agron. J.* 100, 760–764 doi: 10.2134/agronj2006.0282
- Sanjana Reddy, P. (2017). Pearl Millet, *Pennisetum glaucum* (L.) R. Br., in *Millet and Sorghum: Biology and Genetic Improvement*, ed J. V. Patil (West Sussex: Wiley-Blackwell), 49–86. doi: 10.1002/9781119130765.ch2
- Santos, I. G. D., Carneiro, V. Q., Santénna, I. D. C., Cruz, C. D., Carvalho, C. G. P. D., Filho, A. B. B., et al. (2019). Factor analysis and GGE biplot for environmental and genotypic evaluation in sunflower trials. *Funct. Plant Breed. J.* 1, 29–40. doi: 10.35418/2526-4117/v1n2a3
- Tefera, E. (2018). GGE biplot phenotypic stability analysis of soybean [*Glycine max* (L.) Merrill] genotypes. *J. Plant Breed. Crop Sci.* 10, 324–335. doi: 10.5897/JPBSCS2018.0754
- Xie, Y., Xu, J., Tian, G., Xie, L., Xu, B., Liu, K., et al. (2020). Unraveling yield-related traits with QTL analysis and dissection of QTL × environment interaction using a high-density bin map in rapeseed (*Brassica napus* L.). *Euphytica* 216:171. doi: 10.1007/s10681-020-02708-5
- Yadav, O. P., and Rai, K. N. (2013). Genetic improvement of pearl millet in India. *Agric. Res.* 2, 275–292. doi: 10.1007/s40003-013-0089-z
- Yan, W. (2001). GGE biplot—a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agron. J.* 93, 1111–1118 doi: 10.2134/agronj2001.9351111x
- Yan, W. (2002). Singular value partitioning for biplot analysis of multi-environment trial data. *Agron. J.* 4, 990–996. doi: 10.2134/agronj2002.9900
- Yan, W., Hunt, L. A., Sheng, Q., and Szlavics, Z. (2000). Cultivar evaluation and megaenvironment investigation based on the GGE biplot. *Crop Sci.* 40, 597–605. doi: 10.2135/cropsci2000.403597x
- Yan, W., Kang, M. S., Ma, B., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* 47, 643–653. doi: 10.2135/cropsci2006.06.0374
- Yan, W., Pageau, D., Frégeau-Reid, J. A., and Durand, J. (2011). Assessing the representativeness and repeatability of test locations for genotype evaluation. *Crop Sci.* 51, 1603–1610. doi: 10.2135/cropsci2011.01.0016
- Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: principles and applications. *Can. J. Plant Sci.* 86, 623–645 doi: 10.4141/P05-169

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Sanjana Reddy, Satyavathi, Khandelwal, Patil, Gupta, Sharma, Mungra, Singh, Narasimhulu, Bhadarge, Iyanar, Tripathi, Yadav, Bhardwaj, Talwar, Tiwari, Kachole, Sravanti, Shanthi Priya, Athoni, Anuradha, Govindaraj, Nepolean and Tonapi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Sorghum Pan-Genome Explores the Functional Utility for Genomic-Assisted Breeding to Accelerate the Genetic Gain

Pradeep Ruperao¹, Nepolean Thirunavukkarasu², Prasad Gandham¹, Sivasubramani Selvanayagam¹, Mahalingam Govindaraj¹, Baloua Nebie³, Eric Manyasa⁴, Rajeev Gupta¹, Roma Rani Das¹, Damaris A. Odeny⁴, Harish Gandhi¹, David Edwards⁵, Santosh P. Deshpande^{1*} and Abhishek Rathore^{1*}

¹ International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India, ² Genomics and Molecular Breeding Lab, ICAR-Indian Institute of Millets Research, Hyderabad, India, ³ Sorghum Breeding Program, International Crops Research Institute for the Semi-Arid Tropics, Bamako, Mali, ⁴ Sorghum Breeding Program, International Crops Research Institute for the Semi-Arid Tropics, Nairobi, Kenya, ⁵ School of Biological Sciences and Institute of Agriculture, The University of Western Australia, Perth, WA, Australia

OPEN ACCESS

Edited by:

Sean Mayes,
University of Nottingham,
United Kingdom

Reviewed by:

Bernardo Ordas,
Consejo Superior de Investigaciones
Científicas (CSIC), Spain
Fernando Martinez,
Sevilla University, Spain

*Correspondence:

Abhishek Rathore
a.rathore@cgiar.org
Santosh P. Deshpande
s.deshpande@cgiar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 18 February 2021

Accepted: 28 April 2021

Published: 01 June 2021

Citation:

Ruperao P, Thirunavukkarasu N, Gandham P, Selvanayagam S, Govindaraj M, Nebie B, Manyasa E, Gupta R, Das RR, Odeny DA, Gandhi H, Edwards D, Deshpande SP and Rathore A (2021) Sorghum Pan-Genome Explores the Functional Utility for Genomic-Assisted Breeding to Accelerate the Genetic Gain. *Front. Plant Sci.* 12:666342. doi: 10.3389/fpls.2021.666342

Sorghum (*Sorghum bicolor* L.) is a staple food crops in the arid and rainfed production ecologies. Sorghum plays a critical role in resilient farming and is projected as a smart crop to overcome the food and nutritional insecurity in the developing world. The development and characterisation of the sorghum pan-genome will provide insight into genome diversity and functionality, supporting sorghum improvement. We built a sorghum pan-genome using reference genomes as well as 354 genetically diverse sorghum accessions belonging to different races. We explored the structural and functional characteristics of the pan-genome and explain its utility in supporting genetic gain. The newly-developed pan-genome has a total of 35,719 genes, a core genome of 16,821 genes and an average of 32,795 genes in each cultivar. The variable genes are enriched with environment responsive genes and classify the sorghum accessions according to their race. We show that 53% of genes display presence-absence variation, and some of these variable genes are predicted to be functionally associated with drought adaptation traits. Using more than two million SNPs from the pan-genome, association analysis identified 398 SNPs significantly associated with important agronomic traits, of which, 92 were in genes. Drought gene expression analysis identified 1,788 genes that are functionally linked to different conditions, of which 79 were absent from the reference genome assembly. This study provides comprehensive genomic diversity resources in sorghum which can be used in genome assisted crop improvement.

Keywords: sorghum, pan-genome, diversity, SNP, gPAV, GWAS, drought genes

INTRODUCTION

Sorghum (*Sorghum bicolor*) is a multi-utility cereal of global importance, and a major food crop in sub-Saharan Africa and South Asia (Ritter et al., 2007; Motlhaodi et al., 2014). It is typically a diploid species ($2n = 20$) with an estimated genome size of the ~800 Mb sequence (Price et al., 2005). It provides important primary and secondary products, such as food, fodder, starch, fibre,

biofuels, alcohol, dextrose syrup as well as other products. It is domesticated and further bred for diverse use as food, fodder, and bioenergy in different agro-climatic conditions (Li et al., 2010) and shows a wide diversity at the genome level (Kong et al., 2000; Hart et al., 2001).

A draught sorghum genome assembly of 730 Mb was initially prepared for *Sorghum bicolor* Moench (Paterson et al., 2009), followed by an improved assembly of 732.2 Mb, covering ~91.5% of the genome (McCormick et al., 2018). Recently, a sorghum reference genome assembly for the “Rio” line was generated comprising 729 Mb (Cooper et al., 2019). Each of these genome assemblies is limited to its respective accession and does not reflect the diversity of genes in this species.

The presence or absence of genes or genomic regions among genotypes is an important form of genomic variation in plants, and genes can be categorised into core and variable within the species (Saxena et al., 2014; Golicz et al., 2016b). The collection of these core and variable genes is known as pan-genome. Studying the pan-genome from a large number of genotypes enhances the understanding of species diversity, domestication and breeding history, and provides complete characterisation of species genes content diversity as demonstrated in rice (Wang et al., 2018) and tomato (Gao et al., 2019).

Several approaches are available to construct a pan-genome (Golicz et al., 2016b). The classical approach of whole-genome assembly of all genotypes was initially implemented in bacteria, and later developments led to the complementary method to “iteratively map and assemble,” the unmapped sequence reads, demonstrated in *B. oleracea* (Golicz et al., 2016a), *B. napus* (Hurgobin et al., 2018), bread wheat (Montenegro et al., 2017), and pigeon pea (Zhao et al., 2020). The whole genome assembly and comparison method has the advantage in that it can place almost all individual specific genes in a genomic context, but suffers from the inability to distinguish assembly or annotation errors from true biological variation (Bayer et al., 2017). It is also unsuitable for large population studies due to the expense of sequencing, assembling, and comparing large numbers of genomes. In contrast, the iterative assembly approach can cost effectively assess large numbers of genotypes for gene presence/absence variation and hence identify genes that may be relatively rare in a population and not samples in whole genome assembly approaches, though without additional long read data, it is unable to place many of the newly identified genes. Hence the iterative assembly method is most suited for large population diversity studies.

Hence, we assembled a pan-genome using reference and re-sequenced genomes for genetically diverse race-specific sorghum accessions. The sorghum pan-genome was initiated with the reference genome obtained from JGI on Phytozome (McCormick et al., 2018), followed by adding to this reference with novel genome sequences from 176 sorghum accessions.

We provided structural and potential functional aspects of this pan-genome in the form of genes, single nucleotide polymorphism (SNP) and gene presence and absence variations (PAV). The utility of the pan-genome was demonstrated by identifying candidate functional genes using publicly available SNP chip data, genome-wide association studies and

gene-expression assays. These sorghum pan-genome resources will be useful for achieving the sustainable development goals in developing countries by accelerating the genetic gain in arid and semi-arid ecologies.

MATERIALS AND METHODS

Pan-Genome Assembly and Annotation

The pan-genome was assembled using iterative mapping and assembly approaches. The assembly was initiated with a sorghum reference assembly v3.0.1 to map sorghum accessions whole-genome sequence data iteratively. Reads from 176 sorghum accessions with a minimum of 10X coverage sequence data were mapped to the sorghum reference v3.0 (McCormick et al., 2018) using Bowtie2 (Langmead and Salzberg, 2012) v2.3.4, and unmapped reads were assembled with IDBA_UD assembler (Peng et al., 2012) and the assembled contig sequence more than 500 bp length was only considered and appended to reference genome sequence. The resulting final assembly sequence was compared with NCBI non-redundant nucleotide databases using BLASTn and the sequences with homology to sorghum mitochondria (NC_008360), chloroplast (MK348612) also the sequences having homology outside the green plant group Viridiplantae taxonomy group (Taxonomy ID: 33090) were removed. The remaining sequences were self-compared with nucmer search (<http://mummer.sourceforge.net/>) and sequences with >90 percent coverage with greater 90 percent identity were removed to maintain the non-redundancy of the novel sequences. REPEATMASKER-v4.0.7 (Smit et al., 2000) masked repetitive elements using sorghum as the species. The sorghum expressed sequence tags (ESTs) from GenBank were aligned with tBLASTx and genes were predicted using AUGUSTUS v3.3.2, supporting the EST alignments. The gene models having fewer than 300 bp in length were filtered out and the remaining genes supporting either EST alignments or hisat2 (Kim et al., 2019) alignments (RNASeq read from 25 accessions, **Supplementary Table 1**) further used for functional annotation against uniref90 (database downloaded in May 2020).

Gene Presence-Absence Variations (gPAVs)

Whole-genome sequence reads of all 354 sorghum accessions were mapped with Bowtie2 v2.3.4 (Langmead and Salzberg, 2012) to pan-genome assembly with a wide insert size range between 0 and 1,000 bp. The gene PAVs were defined based on sequence reads coverage mapped to respective genes as described by Golicz et al. (2016a). Genes models on contigs longer than 1 Kbp were used in this analysis. PAV converted into the binary matrix and with 1,000 bootstrap resampling were used to estimate the genetic relationship among the accessions with R “ape” package (Paradis et al., 2004) to construct an NJ tree and visualised in iTOL tree viewer (Letunic and Bork, 2019).

The core genes were defined as the genes present in all the accessions, whereas the variable genes are the genes missing in one or more accessions. The *in-house* developed script was used to define the core and variable genes from the PAV matrix. Core and variable genes were compared for gene length, exon number,

synonymous SNPs, non-synonymous SNPs, and Ka/Ks. The mean count for each sample size of core and pan-genes present in all possible combinations of 354 accessions was plotted. The protein sequences of *Zea mays*, *Setaria italica*, *Brachypodium distachyon*, and *Oryza sativa* were downloaded from the public database UniProt for cluster analysis. All protein sequences were compared using all-by-all BLASTp followed by MCL for gene clustering into gene families with default parameters. The gene enrichment analysis was performed with Fisher exact test from R “topGO” package (Alexa et al., 2006) using “Elim” method.

SNP Discovery and Annotation

The sorghum whole genome sequence reads of 354 accessions were quality trimmed using Trimmomatic (Bolger et al., 2014) and mapped to pan-genome using Bowtie2 v2.3.4 (Langmead and Salzberg, 2012) allowing to map paired reads. The aligned reads in SAM format converted to BAM format using samtools (Li et al., 2009) followed by filtering out the read duplication with Picard tools (<http://broadinstitute.github.io/picard>). Variants against the reference (pan-genome) were called with GATK v.4.1 (McKenna et al., 2010) and directed to quality filtered with vcftools v.0.1.13 (Danecek et al., 2011). The variant sites having missing genotypes of more than 0.15 and minor allele count <2 were excluded and the remaining sites were used for downstream analysis such as SNP functionally annotated with SnpEff v.4.3 (Cingolani et al., 2012).

Sorghum Diversity and Population Structure

A subset of 216 diverse sorghum accessions from 354 set with known sorghum race information (Valluru et al., 2019), was used for genetic diversity and population structure assessment. A total of 1.12 million filtered SNPs from sorghum race accessions were retained for downstream analysis. The STRUCTURE v2.3 (Hubisz et al., 2009), was used to estimate the population structure using the admixture model. The tested K was set from 2 to 5 and optimal K for population structure was defined with the structure program. With the same SNP set, PCo analysis was done with R labdsv package (<https://CRAN.R-project.org/package=labdsv>) and phylogeny analysis performed using 1000 replicates with R “ape” package (Paradis et al., 2004) and visualised in iTOL tree viewer (Letunic and Bork, 2019).

Genome-Wide Association Analysis (GWAS)

Two different mapping populations having the phenotypic data of 10 traits were used for the association study.

Pop1

The phenotype and genotype data associated with plant height (PH), dry biomass (DBM), and starch (ST) were adapted from published work (Valluru et al., 2019) for GWAS analysis. A subset of 227 accessions from the 354 WGS set belonged to four major races of sorghum having representation from Africa, Asia, and America was used. The SNPs corresponding to the above-mentioned genotypes were filtered with vcftools and used for GWAS. In 2016, the PH was recorded from 4 to 16 weeks after

planting with an interval of 2 weeks, DBM and ST was measured at harvest.

Pop2

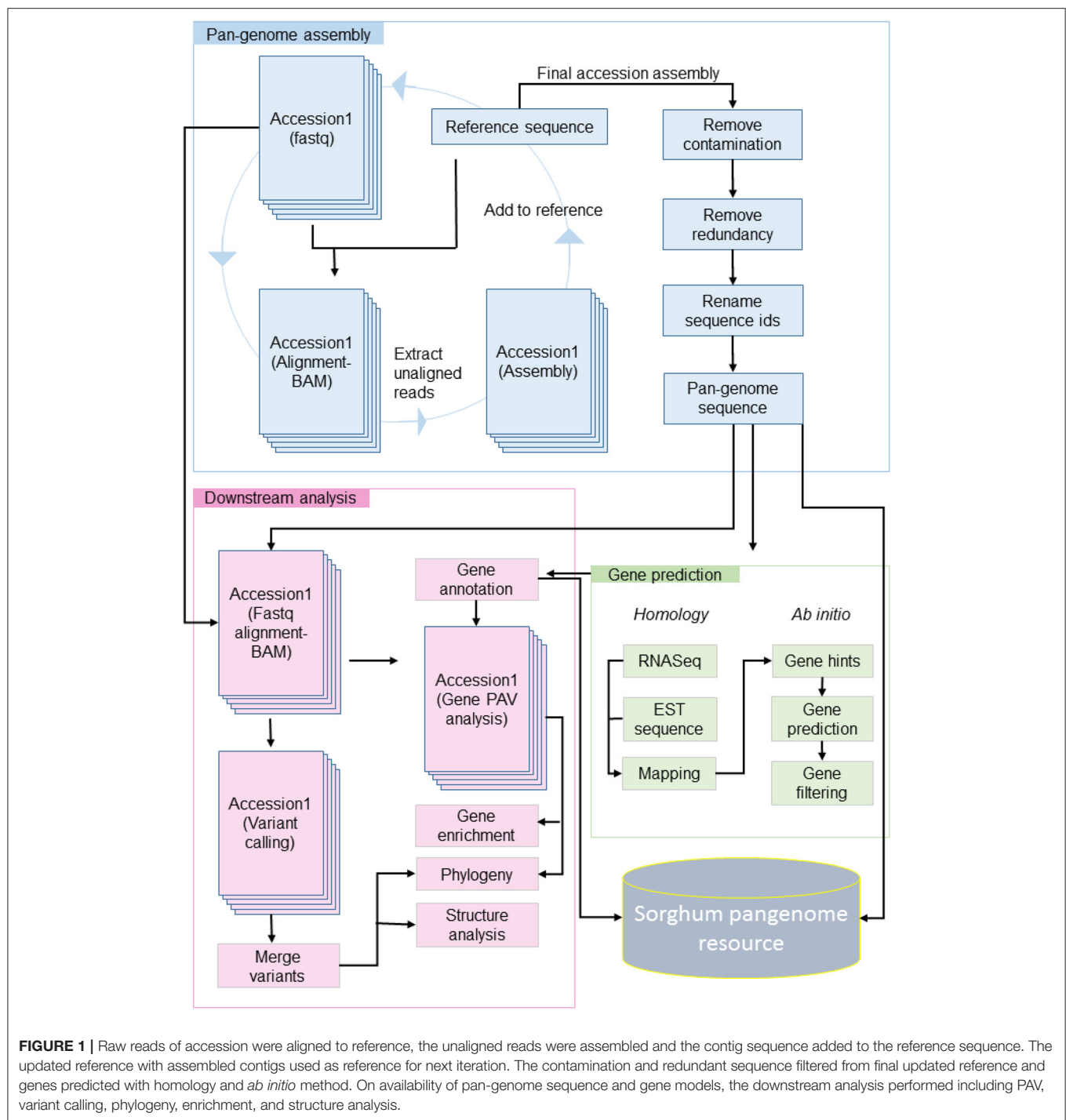
The stay-green fine-mapping population developed by crossing an introgression line cross RSG04008-6 × J2614-11 (Usha Kiranmayee et al., 2020) was used for association study using the pan-genome assembly. The DNA from parents and 152 individuals were isolated and skim-sequenced to produce genotype data to a depth of 0.1X. The sequence reads were QC'd with trimmomatic (Bolger et al., 2014), mapped with bowtie2 (Langmead and Salzberg, 2012) and SNP called with GATK (McKenna et al., 2010) and filtered with vcftools (Danecek et al., 2011) as above said method.

The Pop2 was evaluated with green leaf area (GLA) trait in the *rabi* season of 2012–2013 and 2013–2014 at ICRISAT, Patancheru, India. The GLA percentage was measured from seven to 49 days after flowering (DAF) for every 7 days interval in both years. Additionally, in the year 2013, the phenotypes of glossy (GL), leaf sheath pigment (LSP), plant vigour (V), trichome low (TL), trichome up (TU), soot fly dead hearts (SFDH) traits were recorded in *rabi* (R13), and *kharif* (K13) seasons.

The genotype to phenotype association was performed with GAPIT (Lipka et al., 2012) and the results were initially filtered with Bonferroni cut-off [$-\log_{10}(p\text{-value}) > 2.5$] followed by *p*-value and false discovery rate values <0.05 (close to Benjamini-Hochberg cut-off value) as the significant values. These significant SNPs were further functionally annotated with predicted gene coordinates.

Drought RNASeq Assay Analysis

To demonstrate the utility of the pangenome, we have used a sorghum transcriptome experiment on drought response (Abdel-Ghany et al., 2020) available in the Sequence Read Archive (SRP227627). In this study, the RNASeq data were derived from contrasting genotypes- drought resistant [BT × 623 (DR1) & SC56 (DR2)] and drought susceptible [T × 7000 (DS1) and PI482662 (DS2)] at the seedling stage was obtained. The quality cheque was performed on raw sequence reads using FastQC (Andrews, 2015) followed by cleaning the low-quality reads and removing sequencing adaptors using the Trimmomatic (Bolger et al., 2014) tool. Trimmed reads were aligned to the Sorghum pan-genome using TopHat2 (Kim et al., 2013) and bam files were filtered to remove reads aligned to multiple locations. Differential gene expression was performed on different conditions using Cuffdiff (Trapnell et al., 2010) to compute logFC and *q*-values across all accessions at different conditions (control and treated). A total of eight conditions were analysed to find drought-induced genes after 1 and 6 h of post-treatment (20% PEG treatment). Two biological replicates were analysed for each condition resulting in 32 samples (4 genotypes × 2 conditions × 2-time points × 2 replicates). The differentially expressed genes (DEGs) were determined if the *q* < 0.05 and log2FC is <−2 or >2 ratios between control and treatment for each time point and in each genotype.



RESULTS

Pan-Genome Assembly

Genome sequence data with minimum 10X coverage from earlier studies (Guo et al., 2019; Valluru et al., 2019) were used for pan-genome assembly (**Supplementary Table 1**). The pan-genome was constructed using 176 sorghum accessions using an iterative mapping and assembly approach, similar to Brassica (Golicz et al., 2016a) and pigeon pea (Zhao et al., 2020) (**Figure 1**).

On an average, each iteration of the process added 1.9 Mb of sequence to the reference (**Supplementary Figure 1**) and a total of 263.7 Mbp was assembled. Of these, 89.2 Mb of the sequence were removed as contaminants (including chloroplast and mitochondrial sequences) and/or duplicated contigs. The final resulting pan-genome contained 210,805 contigs with a total length of 883.3 Mb (**Figure 2**) with a minimum contig size of 500 bp. Gene density on the contigs added by this pan-genome exercise was lower than on assembled chromosomes

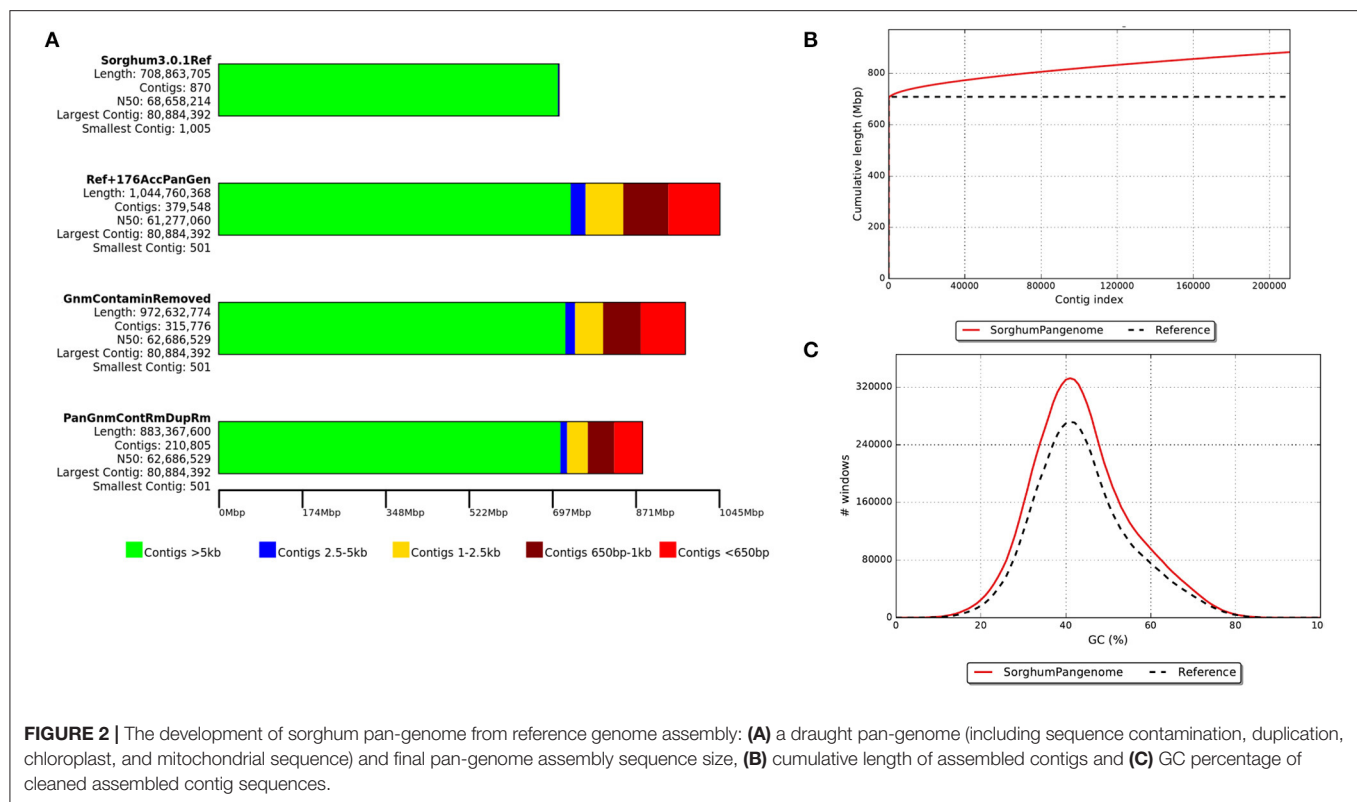


FIGURE 2 | The development of sorghum pan-genome from reference genome assembly: **(A)** a draught pan-genome (including sequence contamination, duplication, chloroplast, and mitochondrial sequence) and final pan-genome assembly sequence size, **(B)** cumulative length of assembled contigs and **(C)** GC percentage of cleaned assembled contig sequences.

but comparable to the density observed on the reference unplaced scaffolds (Figure 3).

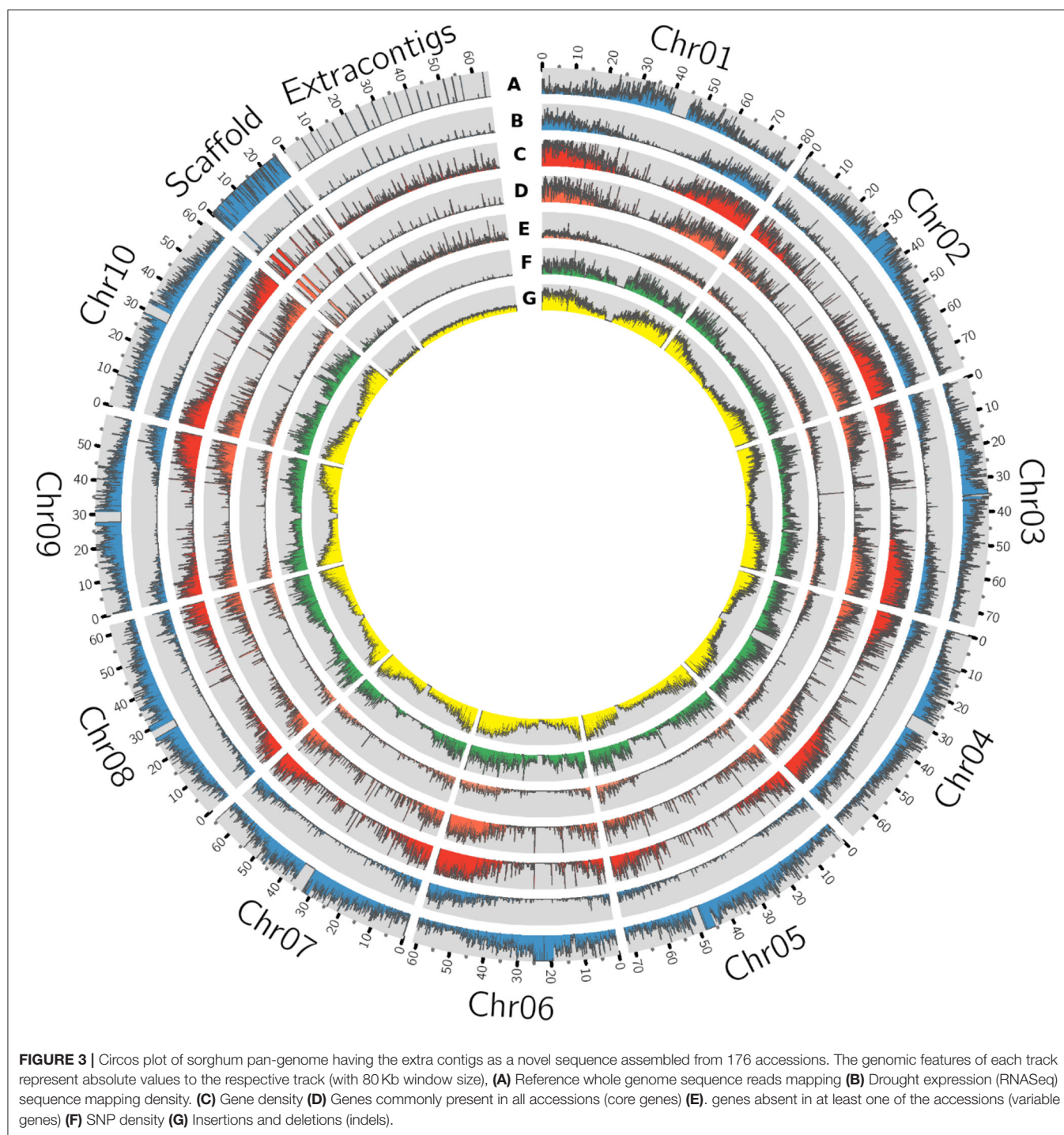
The pan-genome showed an increase of 24.6% (174.5 Mb) over the reference genome, which was the second-biggest increase of any previously reported pan-genome after the tomato pan-genome. The increase in tomato pan-genome size was captured a 42% non-reference sequence from 725 accessions including the wild relatives (Gao et al., 2019). In other species, an increase in sequence size of 3.3% in wheat (Montenegro et al., 2017), 4% in *Oryza sativa japonica*, 6% in *Oryza sativa indica* (Yao et al., 2015), 5% in *Brachypodium distachyon*, and 20% in *Brassica oleracea* (Golicz et al., 2016a) was documented. The relative increase in sorghum pan-genome assembly size indicated that the presence of high level of genome diversity contributed by the accessions used in this study.

The assembled sequence was annotated using a strategy called combining evidence-based *ab initio* gene prediction. RNASeq (Guo et al., 2019) mapping hints from the 25 accessions used for *ab initio* gene prediction and the 3,589 genes supporting the mapped expressed sequence tags (EST) sequences were retained. We identified 11,057 to 17,616 variable genes in the 176 genomes, with an average gene sequence length and exons per gene of 1,567 bp and 3.6, respectively. The gene length and exons in core genes were more than the variable genes comparatively (Figure 4).

Sorghum Pan-Genome Gene PAV (gPAV)

The gPAV in genes among the sorghum accessions could reveal the genetic changes that can be used to infer the phylogenetic

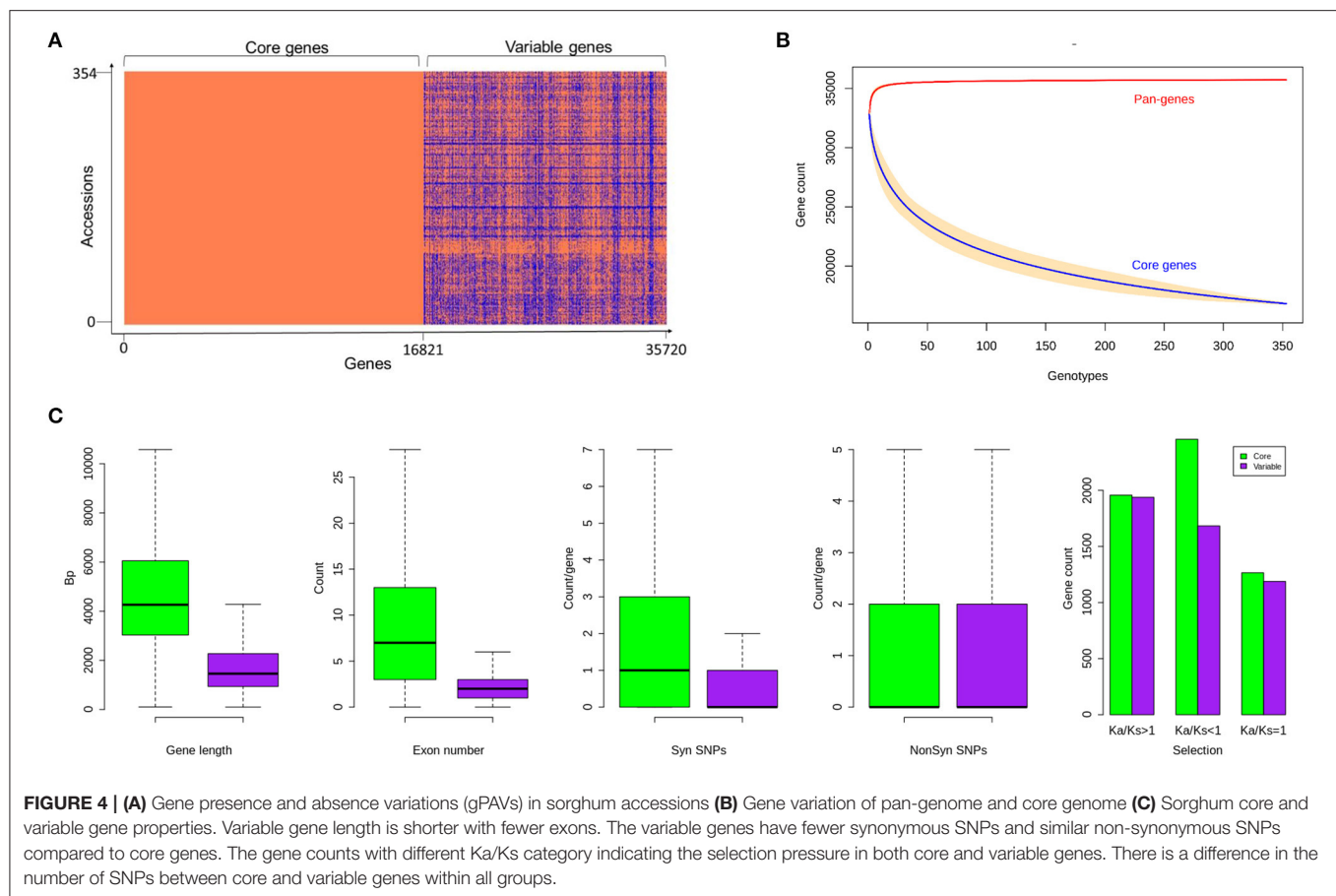
history as well as to select the potential targets for breeding. To identify the gPAVs, sequence reads were mapped to the pan-genome contigs and genes were scored as present or absent based on the mapped sequence read coverage (Supplementary Table 2, Figure 5). For a given gene, to assess the gene loss event, the mapping of the whole genome sequence reads was measured. On an average, each sorghum accession contained 32,795 genes (Supplementary Table 3), of which 16,821 (47%) were core genes or in other words, they were shared by all remaining accessions. Comparatively, tomato (Gao et al., 2019) (74.2%), maize (Hirsch et al., 2014) (39%), *Arabidopsis thaliana* (Contreras-Moreira et al., 2017) (70%), wheat (Montenegro et al., 2017) (64%), pigeon pea (Zhao et al., 2020) (86%), *Brassica rapa* (Lin et al., 2014) (87%), *O. sativa* (Schatz et al., 2014) (92%), and, *Brassica napus* (Hurgobin et al., 2018) (62%) had higher number of genes (Bayer et al., 2020). On the other hand, 18,898 genes were variable/accessory genes (Figure 4A), of which 30 genes were uniquely present (indicating that the genes are present in any one accession but absent in remaining all accessions) and 3,183 (8.9%) were uniquely absent (indicating that the genes present in all accessions but absent in any one accession) (Figure 5). Variable genes were found shorter and had fewer exons per gene when compared to core genes (Figure 4C) which were in agreement with *O. sativa* and *A. thaliana* crop studies (Bush et al., 2014; Schatz et al., 2014; Golicz et al., 2016b). Based on gPAVs from 354 cultivars, we estimated the sorghum pan-genome had a closed type of pan-genome (Figure 4B), with 30 genes were uniquely present and 3,183 genes were uniquely absent. The uniquely present genes were fewer than



the wheat (49 unique genes per cultivar) (Montenegro et al., 2017) and *B. oleracea* (37 unique genes per cultivar) (Golitz et al., 2016a). Of the 30 genes uniquely present in any single sorghum accession, nine such genes were reported from Macia accession alone (Figure 5). Extending the population size and including the wild relatives could further increase the measure of the gene content of this species (Figure 4B) (Golitz et al., 2016b).

Gene Functional Analysis

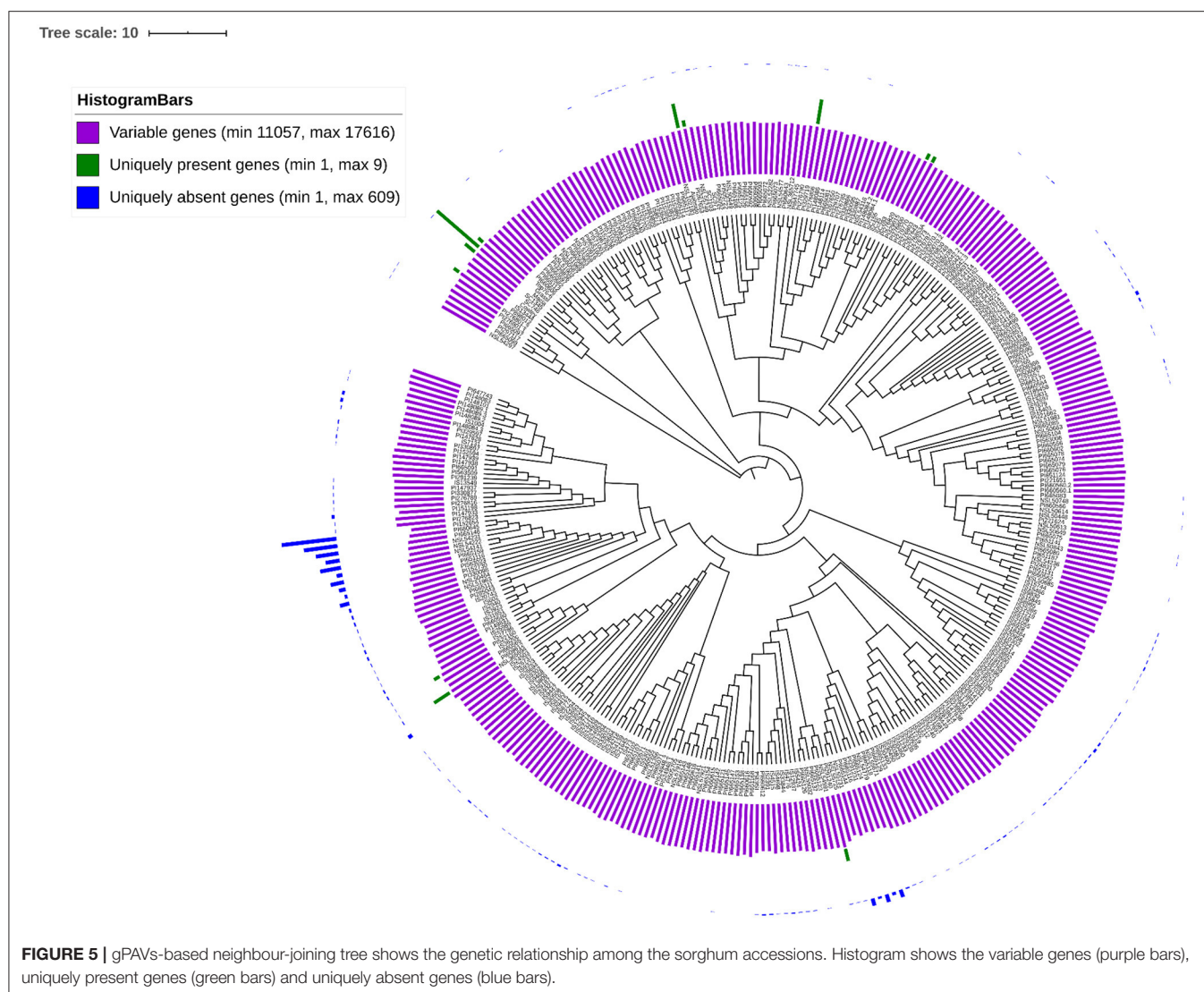
We identified enriched biological pathways by performing gene enrichment analysis using the R topGO package. The significantly enriched pathways related to responsive genes were identified (Figure 6). A total of 94 most significantly enriched genes (Supplementary Table 4) for biological process pathways are shown in Supplementary Figure 2. The gene ontology (GO) enrichment analysis showed that the genes



were enriched in response to chemical, hormone, organic substance, stress, auxin and abiotic stimulus (**Figure 6A**). It was noted that most of the pathways were related to plant response to stimulus and chemicals. The gene enrichment among stress responses genes including water deprivation (GO:0009414), desiccation (GO:0009269), abiotic stimulus (GO:0009628), chemical stimulus (GO:0042221), and stress (GO:0006950) were reported in a reference set of genes (Woldeamayrat and Ntwasa, 2018). The gPAV-based enrichment on assembled genes from the sorghum pan-genome has added the response of the genes to auxin (GO:0009733), hormone (GO:0009725), organic substance (GO:0010033), hypoxia (GO:0001666), and decreased oxygen levels (GO:0036294). The functional annotation of the variable genes highlighted the genes involved in response to biotic and abiotic stress indicating the possible evolutionary adaptive traits (Lasky et al., 2015). Macia (9 genes), Ajabsido (4 genes), and PI329719 (4 genes) were identified with a number of trait-specific genes (**Figure 5**), which could be used as potential donors for trait improvement. It was observed that the above-mentioned unique genes were involved in response to the stimulus (GO:0050896), chemical (g8132, GO:0042221), and arsenic-containing substance (g24192, GO:0046685) (**Supplementary Table 5**).

Variant Discovery

We identified a large number of variants (single nucleotide polymorphism (SNPs) and indels) by mapping the sorghum population whole-genome sequence reads to sorghum pan-genome assembly using GATK. Of the total of 2.0 million SNPs, 91,319 were in the extra contig (assembly) sequence (**Supplementary Table 6**). The SNP density in extra contigs (0.52/Kbp) was much less compared to the density in the reference genome assembly (2.72/Kbp) (**Figure 3**). The SNP annotation results illustrated the highest number of SNPs in intergenic region (40%) followed by upstream (22.5%), downstream (21.4%), intron (8.8%), and exon (3.6%) regions with an overall Ts/Tv ratio of 1.92. Chromosome 4 had the highest number of SNPs (251,830), followed by chromosomes 1, 2, 3, 5, 10, 8, 6, 9, and 7, respectively. Chromosome 7 had the fewest number of SNPs (119,019) with the highest density of 0.55/Kbp and chromosome 4 had the least density of 0.27/Kbp (**Supplementary Table 6**) (**Figure 3**). The presence of more SNPs and indels in the telomeres compared to centromeres explained the higher gene activity toward the telomeres supporting the SNPs and indels density reported in historically important grain sorghum genotypes (BTx623, BTx642, and Tx7000) (Evans et al., 2013) (**Figure 3**). The SNP annotation reported the frequency of synonymous SNPs in the core genes was much higher than



in the variable genes (**Figure 2**). This was in contrast to the higher mis-sense SNPs in core pigeon pea genes to variable genes early reported (Zhao et al., 2020). We detected genome-wide indels of various size (**Supplementary Figure 3A**) and the genes featuring indels has reduced proportionally to the size of the indels (**Supplementary Figure 3B**). On increasing the indel size, the number of the indels decreases in both gene and genome-wide sequence. The overall indels count from the sorghum accessions used in this study was much higher than the indels earlier reported in the six sorghum accessions (Yan et al., 2018). A total of 36,097 genes had 983,060 CNVs among the sorghum accessions used in this study. The K_a/K_s ratio estimating the balance between neutral mutations, purifying selection, and beneficial mutations on a set of core and variable genes exhibited that, core gene count under positive selection were significantly close to variable gene count compared to genes under the negative selection pressure (**Figure 4**).

The maximum (432,286) and minimum (2,854) number of SNPs were identified in sorghum accessions PI267614- NSL54318 and IS3693- IS23514, respectively (**Supplementary Table 7**). The accessions NSL54318 (849,052) and IS3693 (17,084) had the maximum and minimum polymorphism, respectively, with sorghum pan-genome assembly sequence (**Supplementary Table 8**). The SNPs were validated with 3K SNPs Infinium array (Bekele et al., 2013) and of the 2,980 mapped flanking SNPs sequence, only 20 did not map on pan-genome assembly. The overall alignment rate was 99.33%, from the mapped 2,980 SNPs array sequences (**Supplementary Table 9, Figure 6B**). Among them, 37 SNP sequences were mapped to extra contigs (novel sequence assembly) and 150 (5%) did not represent any GATK SNPs calls (29 SNPs from extra-contigs). In addition to the core SNPs of the array sequence, more SNPs were identified in the flanking sequence. Out of 15,383 GATK SNPs on the mapped array sequence, 15,314 SNPs were validated with the GATK called

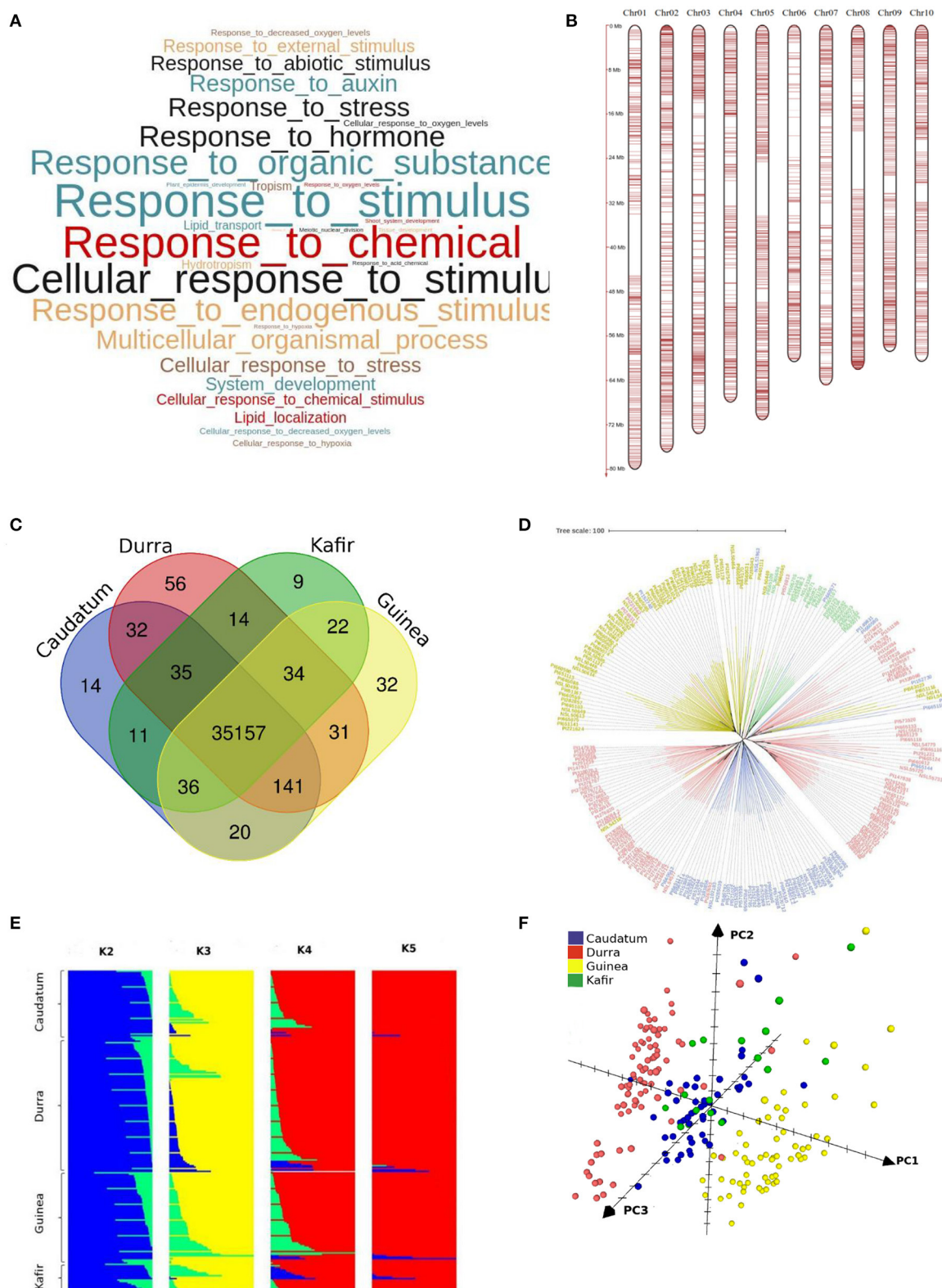


FIGURE 6 | (A) Significantly enriched top GO terms among the variable genes ($p < 0.05$) **(B)** distribution of Infinium SNP array markers on chromosomes **(C)** specific and common genes across races **(D)** neighbour-joining tree shows the genetic relation among the sorghum accessions belonged to different races (blue-Caudatum, red-Durra, green-Kafir and yellow-Guinea) **(E)** structure analysis of sorghum population with K2 to K5 and **(F)** sorghum accessions grouped by caudatum, durra, guinea and kafir race through principal co-ordinate analysis (PCo).

allele (**Supplementary Table 9**). Finally, on validation with array SNPs, the overall GATK SNP calling reported 99.9% accuracy.

To understand the genetic relationship of the 354 sorghum accessions, a neighbour joining (NJ) tree was constructed with the SNPs (**Supplementary Figure 4**). The accessions were arranged in many sub-groups indicating the possible sorghum race accessions. To assess the race-specific accessions, the 216 known sorghum race accession bootstrapped to construct an NJ tree. The NJ tree showed the subgroups of sorghum accessions according to the races except a few, indicating the hybridisation process in the past (**Figure 6D**). For example, PI221662, a durra race accession was genetically related to the guinea race. Similarly, to understand the gene PAV-based genetic relation, the phylogenetic relationship among sorghum accessions was assessed by distance-based 1,000 bootstrap replicates and represented through the NJ tree. Among the 35,719 total genes, 53% exhibited the genic variations to estimate the relationship among the accessions (**Figure 5**). The largest number of genes uniquely present and absent genes was found in Macia (9 genes) and PI660645 (372 genes), which indicated the evolutionary distance from other accessions.

With the known four races (Obilana et al., 1996), the structure analysis with the variants set showed the presence of three sub-population (**Figure 6E**), resulting an expected admixture between *caudatum* and *kafir* accessions which was in agreement with early study (Valluru et al., 2019) (**Figure 4**). The result was also validated by the PCo, where *durra* and *guinea* sorghum races displayed identifiable clusters, because of the available sequence representation through pan-genome, while *caudatum* and *kafir* accessions exhibited the admixers (**Figure 6F**). The earlier principal component study shows the mixed grouping of *guinea* and *kafir* accession in the (Sapkota et al., 2020), indicating the missing sequence representation for all race in the single reference genome.

Variation of Sorghum Race-Variable Genome

Sorghum pan-genome analysis has identified 18,898 variable genes, and the gene cluster analysis identified 11,470 gene families, of which un-clustered genes (6,057) included 556 from the non-reference genes and the remaining 5,501 were reference genes. Among these un-clustered genes, 3,195 were orthologous to *Zea mays*, *Setaria italica*, *Brachypodium distachyon*, and *Oryza sativa* and the remaining 2,862 were paralogous. Among the total variable genes, a total of 111 genes are race-specific and the gene shares among four sorghum races showed that the *durra* and *guinea* had a maximum of 56 and 32 unique genes, respectively, making them more diverse than the other two races with 14 (*caudatum*) and 9 (*kafir*) unique genes (**Figure 6C**). The gene annotations suggested that the unique genes from *durra* were associated with heat shock protein, LRR repeat protein, L-type lactin-domain receptor, ABC transporter family proteins, and Ras-related proteins. *Guinea* group had the unique genes associated with disease resistance protein, beta-glucosidase proteins, NRT1/PTE protein family, and Alpha/beta-Hydrolases superfamily proteins (**Supplementary Table 10**). The

gene uniqueness to specific races possibly reflected the selection of the genotypes for adapting to the respective ecological conditions (Upadhyaya et al., 2017).

GWAS

Two populations namely, Pop1 (Valluru et al., 2019) and Pop2 (Usha Kiranmayee et al., 2020) were used for GWAS to understand the functional utility of the pan-genome. Pop1 had 216 accessions with the phenotypes of DBM, PH, and ST while Pop2, a stay-green fine-mapping population with 190 segregates, had green leaf area (GLA), GL, V, LSP, SFDH, TL, and TU.

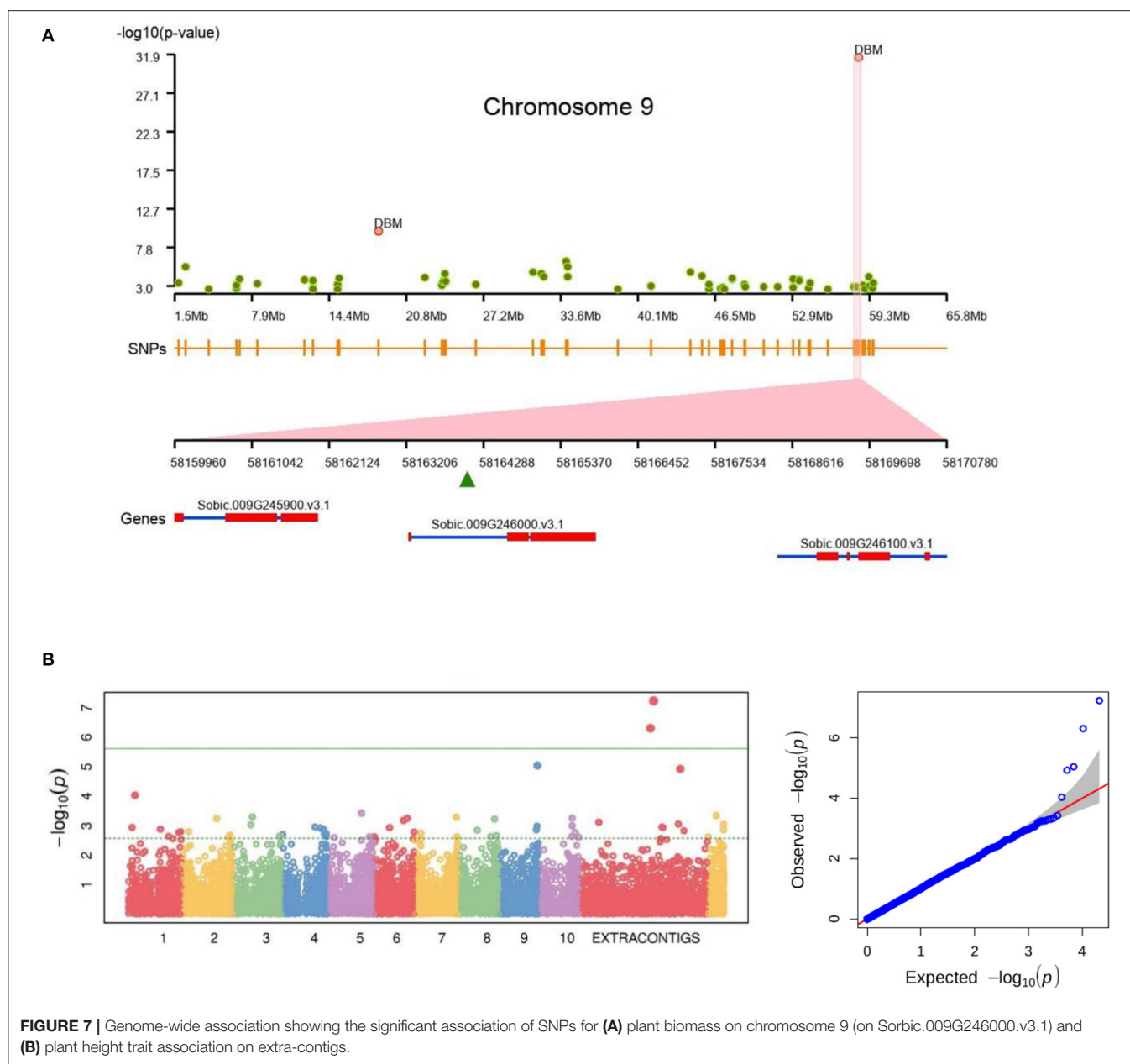
In Pop1, the SNPs were further filtered by accessions and on applying the SNP quality philtres, which retained 1.12 million SNPs for association analysis. Pop2 having sequence data of 190 genotypes processed to map to pan-genome and 109,338 SNPs were used for GWAS.

We identified a total of 397 unique SNPs having significant association (having *p*-value and false discovery rate below 0.05) in both Pop1 and Pop2 traits, of which 216 SNPs were commonly mapped with multiple traits. Most of these SNPs distributed on chromosome 10 (120 SNPs) followed by chromosome 6 (69 SNPs). The reference genome alone had 385 SNPs and the rest of the SNP-trait associations located on the unmapped read sequence assembly.

For the Pop1, a total of 36 SNPs had a significant association across three traits (**Figure 7A**) (**Supplementary Table 11**). Among them, seven were located on newly assembled contigs (DBM and PH) (**Figure 7B**), three were from unplaced reference contigs and the remaining 26 are from chromosome sequence. Among the 36 linked SNPs, 10 were genic and the remaining 26 were inter-genic regions (**Supplementary Table 12**). Three of the genic SNPs were associated with DBM while six were associated with PH and the remaining one co-mapped to both DBM and PH. From the 10 associated genes, three genes (Sobic.002G022500, Sobic.003G173400, and Sobic.004G350800) were from the core gene set and the remaining belonged to variable genes.

From the Pop2, the GLA, at various stages (**Supplementary Table 13**), associated with 219 SNPs including 111 genic, distributed across all chromosomes including the pan-genome assembly contigs (**Supplementary Table 11**). GL, LSP, SFDH, TL, and TU traits were associated with 129 and 103 significant SNPs in *Rabi* (R13) and *Kharif* (K13) seasons, respectively (**Supplementary Table 11**). The majority of the SNPs were associated with chromosome 10 followed by 5 and 6 in both the seasons. Among them, a total of 96 SNPs was mapped across seasons and a total of 18 and 196 showed season-specific association in K13 and R13, respectively. Interestingly, only four genic SNPs were associated with TU in K13, whereas 63 were associated in R13 explained the season-specific gene regulations. Similarly, SFDH had no association in K13 but had 56 genic SNPs in R13 season (**Supplementary Table 11**).

The number of SNPs associated with DBM, PH, ST (Pop1), plant vigour (V), GL, LSP, SFDH, TL and TU, and GLA (Pop2) was 10, 25, 1, 1, 23, 31, 84, 169, 98, and 397, respectively. Among the chromosomes, as many as 392 of the SNPs were



associated with chromosome 10 and only 8 SNPs were associated with chromosome 3 (3 SNPs on scaffolds). The pan-genome assembly contigs hold 15 trait-associated SNPs, an additional genetic resource for the sorghum breeding program.

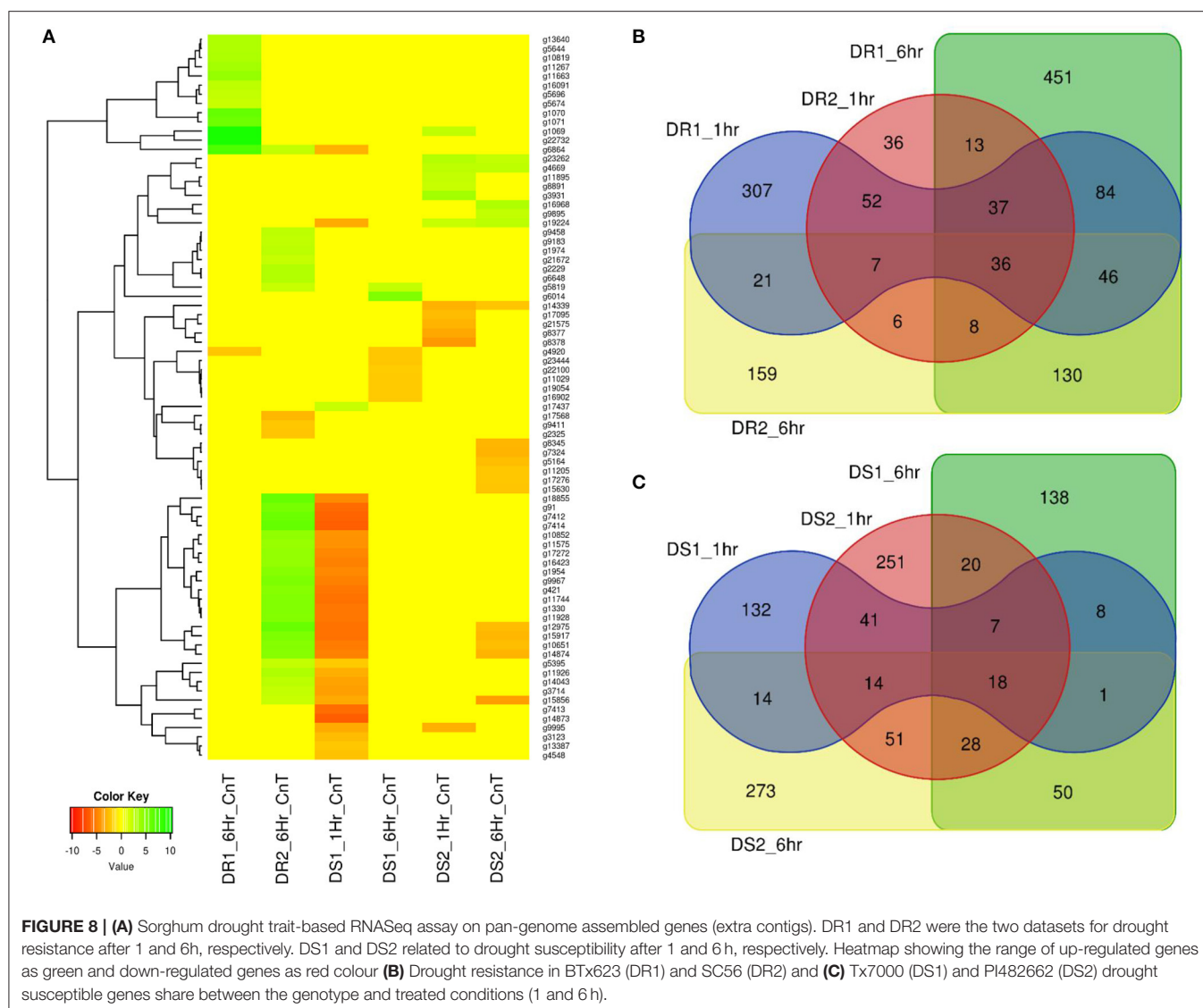
Of the total 183 GWAS SNPs directly associated with gene functions, the DBM and PH (from Pop1) were associated with 10 genes (off these, 1 gene assembled in this study). In Pop2, 173 genes were distributed as 96, 11, 13, 46, 48 and 1 for GLA, GL, LSP, SFDH, TL, and TU, respectively.

Identification of the Drought Candidate Genes

A sorghum RNASeq data generated from drought-resistant [BTx623 (DR1) & SC56 (DR2)] and susceptible [Tx7000 (DS1)

and PI482662 (DS2)] genotypes at different seeding stages (Abdel-Ghany et al., 2020) were re-analysed and mapped through the newly developed pan-genome. A total of 1,788 genes were significantly affected by drought stress (**Figure 3**) and among them, 79 genes were reported from genes on assembly sequence (extra- contig) (**Figure 8A**) (**Supplementary Table 14**).

The drought-resistance (DR1 and DR2) and drought-susceptibility (DS1 & DS2) samples were phenotyped at two conditions (1 and 6 h). The DR1 and DR2 samples reported (1 h treatment) a total of 590 (450 up and 140 down-regulated) and 195 (180 up- and 15 down-regulated) expressed genes, respectively (**Figure 8B**). Of these, none of the genes reported from the novel sequences, indicating both (DR1 and DR2) were closely related. Additionally, DR1 and the reference sequence



belong to the same genotype and this supports the absence of gene expression from the novel sequence at this condition. When the treatment was extended for 6 h, 14 (13 up and 1 down-regulated) and 34 (31 up and 3 down-regulated) genes from novel sequence were expressed for DR1 and DR2 data-sets, respectively.

Similarly, DS1 and DS2 samples showed (for 1 h treatment) 235 (123 up and 112 down-regulated) and 430 (388 up and 42 down-regulated) genes, respectively (**Figure 8C**). Of these, DS1 and DS2 had 32 (1 up and 31 down-regulated) and 13 (7 up and 6 down-regulated) expressed genes from novel sequence, respectively. After 6 h of treatment, DS1 and DS2 samples reported 270 and 449 expressed genes respectively. Of these, DS1 and DS2 reported 8 (2 up and 6 down-regulated) and 17 (5 up and 12 down-regulated) expressed genes were from the novel sequence, respectively.

Over-all, five drought-related genes were co-mapped with the trait-associated genes. Among the five genes, three traits-linked genes Sobic.001G363200 (GLA), Sobic.007G180300

(GL), and Sobic.010G231900 (TL, TU, and SFDH) were commonly expressed in drought resistance and susceptibility conditions. The remaining two drought resistance specific genes Sobic.005G069800 and Sobic.006G127800 were linked to PH and LSP traits.

DISCUSSION

We built a sorghum pan-genome with an iterative mapping and assembly approaches with 176 of 354 whole genomes sequenced accessions having coverage of more than 10X. The total size of the pan-genome has become 883 Mbp, with a 20% increase (175 Mbp) compare to the reference assembly of 708 Mbp. This level of novel sequence increase probably due to the high level of genetic diversity observed in the respective species (Cuevas and Prom, 2020).

We have generated the pan-genome genomic resource from the diverse sorghum accessions including the basic

and intermediate sorghum races [bicolor (B), *caudatum* (C), *durra* (D), *kafir* (K), and *guinea* (G)]. Comparison of the wide range of sorghum whole genome sequence datasets has enabled to assemble many coding genes that were absent in published sorghum reference genome sequences. The mapping of RNASeq read from 25 accessions on the assembled contigs supports the predicted genes on the novel sequence (**Supplementary Figure 5**) is an additional genetic resource that will enhance the identification of the quantitative trait locus (QTL) and genome-wide association studies (Chen et al., 2014; Yano et al., 2016; Zhao et al., 2020). The earlier pan-genome studies found that non-reference genes have significantly involved in agronomic traits mainly in plant defence responses (Hirsch et al., 2014; Golicz et al., 2016a; Montenegro et al., 2017; Dolatabadian et al., 2020). Similar to the sorghum genes, *B. oleracea* pan-genome genes also showed that nearly 30 percent of reference genes exhibited the gPAV (Golicz et al., 2016a). It is understood that, as the number of genotypes increases, the size of core genes decreases with a relative increase of variable genes (**Figure 4B**). With the 10 sample sizes, *Brassica oleracea* pan-genome had 20% of PAV genes (Golicz et al., 2016a) which was in consistent with simulation with similar population size in *B. distachyon* pan-genome (Gordon et al., 2017). Similarly, pan-genome from 15 Medicago genomes had 42% of sequences share with few accessions (Zhou et al., 2017), which was comparatively similar to the 49% of sorghum variable pan-genes in this study.

The result of the structure groupings correlated with the PCo showing three different clusters with one of them having two groups (*caudatum* and *kafir*). Among the four basic sorghum races used in this study, PCo displayed, *guinea*, and *durra* remain as distinct clusters while *caudatum* with *kafir* classified with mixed genotypes, which is considered as the stable hybrid race among the 10 possible stable combinations of sorghum races (Obilana et al., 1996). Similarly, mixed PCo clusters were also reported earlier with five basic sorghum races, where the sorghum B race was not well-supported genetically and a majority of them share membership with the remaining four genetic groups (Brown et al., 2011).

The genomic features helped the races to group into different clusters. By looking at the race-specific genomic data, each race had distinctive features. The *guinea* group had 37 accessions with race specific genes present in range of 2–13 genes per accessions, whereas *durra* had 2–12 genes in 92 accessions, *kafir* had 2–5 genes in 12 accession, and *caudatum* had 2–4 genes in 15 accessions. The two groups including *durra* and *guinea* were having 56 and 32 distinct genes, respectively, unique to these groups, whereas *caudatum* and *kafir* have on 14 and 9 distinct unique genes, as these groups have the admixture accessions which share genes between the groups.

The functional analysis of variable genes was enriched with GO terms associated with response to light, flower development, salt stress, water, heat, desiccation, temperature, osmotic stress, lipid, gibberellin, and stress. The results supports the earlier gene function based clustering and enrichment analysis exhibiting the similar stress response genes reported in sorghum (Woldeamayrat and Ntwasa, 2018). The plant hypersensitive response annotation in the variable gene was

reported in plant pan-genome analysis (Golicz et al., 2016a; Montenegro et al., 2017; Hurgobin et al., 2018; Zhao et al., 2020).

The development and application of sorghum SNPs have limited to reference genome assembly sequence used in the analysis. The 1.8 million SNP reported earlier on Rio with respect to BTx623 (Cooper et al., 2019), were limited to the single reference genome. Using the whole genome sequence data from 354 sorghum diverse accessions, we identified two million SNPs and 3.9 million indel sites, which represented the functional genome diversity. The density of genetic variation in the novel assembled sequence was low compared to the reference sequence. The reference genome carried most of the conserved essential genes, indicating that the variable sequence has low diversity (**Figure 3**), as reported in the six sorghum accessions from common geographical regions (Yan et al., 2018). The fewer number of SNPs on variable sequence mainly contained genes involved in response to various stress (biotic and abiotic stress tolerance), this finding is well-aligned with the SNPs from disease resistance R genes differentiating sweet and grain sorghum accessions (Zheng et al., 2011). A reference sequence within the pan-genome assembly alone accounted for 95.4% of SNPs and the added assembly sequence from the sorghum population had 4.5% additional SNPs. A total of 2,980 array SNPs from (Bekele et al., 2013) were identified as similar to GATK called (reference-based variant calling) SNPs with 99.33% of true SNPs. The GATK called sorghum SNPs validation rate with array SNPs was higher (99.33%) compared to the non-reference based variant calling methods, for example, the wheat pan-genome SNPs were called with 96.3% accuracy (Montenegro et al., 2017). The abundance of SNPs depends on factors such as mutation events and genome diversity and the SNPs identified in the variable genome can assist in characterising novel metabolic pathways.

Phylogenetic analysis of 354 sorghum accessions using SNPs on the pan-genome demonstrated the mixed groups of diverse biomass genotypes (Valluru et al., 2019), domesticated accessions (Guo et al., 2019) and Chibas sorghum breeding program accessions (Jensen et al., 2020). gPAV-based phylogeny showed a group of 15 accessions having uniquely absent genes in a range of 2–509 genes from the biomass genotypes indicating the wider genetic diversity. The five Chibas sorghum breeding lines (Macia, Ajabsido, SC1345, P898012, and Grassl) had the most unique genes followed by seven domesticated accessions distributed across the phylogenetic tree. On assessing the known sorghum race genotypes from Valluru et al. (2019) phylogeny showed a cluster for each sorghum race. Few accessions of *caudatum* and *guinea* were mixed with *durra* cluster indicating that these are the *caudatum-durra* (CD) and *guinea-durra* (GD) hybrid individuals. Similarly, few accessions were not placed in respective race groups, for example, PI248317 accession was a *durra* race accession placed in *guinea* race cluster which shared the genetic similarity with *guinea* race as DG hybrid individual.

The GWAS performed in the earlier study was limited to the phenotype association only with limited SNPs on the reference genome used (Morris et al., 2013; Kimani et al., 2020). The SNP calling on sorghum pan-genome has enabled the identification of the variants also from non-reference sequence assembly from the genetically diverse accessions. A total of 91,339 SNPs reported

from the assembled sequence were the additional markers used for GWAS. A total of 36 SNPs (from Pop1) were associated with target traits. Among them, 10, 25 and 1 were from assembly sequence (extra contigs) were associated with DBM, PH, and ST, respectively. Additionally, the GLA (from Pop2) had a significant association with five SNPs on extra-contigs. The GLA phenotypes in 2013 and 2014 after 7, 14, 21, 28, 35, 42, 49 days after flowering (DAF) in *rabi* were associated with 219 SNPs. Most of the SNPs were linked with the GLA recorded at the early stage of 7 (linked with 150 SNPs) to 14 DAF (linked with 161 SNPs) (**Supplementary Table 11**). From the flowering stage to 14 days of post-flowering, the GLA expression was significantly linked with 101 common SNPs (two SNPs reported from Extra-Contig101123 at 855 positions and Extra-Contig170379 at 501 base position) (**Supplementary Table 11**). For the phenotypes of GL, LSP, SFDH, TL, and TU in both *rabi* and *kharif* seasons, a total of 147 SNPs were identified, of which 85 were co-mapped in both seasons, 44 were unique to *rabi* and 18 were unique to *kharif*. Out of total 397 associated SNPs, 12 SNPs from novel sequences having significant trait association is an additional gain from the pan-genome assembly.

Most of the associated SNPs linked to genes including NAC-domain protein controls the flowering time and stress response, BTB domain for protein-protein interaction, PSII protein complex for oxygenic photosynthesis, AAI domain protein for lipid transfer protein (LTP). The genes are transcription factors (TFs) such as nuclear TF, reverse transcriptase Ty1/Copia-type domain and BZIP. The genes also associated with ubiquitination pathway proteins such as B-box, F-box, U-box, RING-type, and RING-type E3 ubiquitin transferase protein supporting the sorghumFDB gene family classifications (Tian et al., 2016).

We found 1,788 drought-responsive genes with different seeding stage sequence data mapping on pan-genome assembly, whereas weekly sampled the growing plants and mapping the RNASeq data to reference alone reported the 44% of genes exhibiting the response to drought stress (Varoquaux et al., 2019).

This difference in drought expression was expected between the seedling samples (in 1–6 h difference) compared to root and leaf large scale sampling in 2–17 weeks of pre and post-flowering drought responses (Varoquaux et al., 2019). Similar drought stress gene expression changes were seen in laboratory and greenhouse studies in sorghum genotypes (Johnson et al., 2014; Fracasso et al., 2016). Identifying 79 drought-linked differentially expressed genes on assembly sequence are the additional genes added from this study (**Supplementary Table 15**). These additional genes through pan-genome were mainly involved in the cell membrane, catalytic activity, molecular function regulation, response to the stimulus, metabolic process, cellular, and biological regulation.

The sorghum pan-genome assembly, genes with its annotations, SNPs data sets are available at the sorghum website (<https://doi.org/10.21421/D2/RIO2QM>).

CONCLUSIONS

We constructed and characterised the sorghum pan-genome using the reference genome assembly and the whole-genome

sequence reads of genetically diverse sorghum accessions. The pan-genome had 35,719 predicted genes, which were categorised as core, conserved genes, and variable genes as they exhibited presence and absence variation. The variable genes were enriched with genes response to various stresses. The SNP Infinium array result showed 99% of representation on the pan-genome assembly sequence. About two million SNPs were developed through pan-genome which can use for functional downstream research. The pan-genome resources were validated by assessing the genetic diversity of sorghum races, identification of genes from GWAS and RNASeq studies. These newly generated genomic resources could be used in sorghum genetic gain improvement programs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s. The sorghum pan-genome assembly and annotation are available at dataverse. icrisat.org (<https://doi.org/10.21421/D2/RIO2QM>).

AUTHOR CONTRIBUTIONS

PR and AR conceived and designed the project. PR, PG, and SS carried out the analysis. SS managed computational resources and data management. SD provided sorghum trait data. PR, NT, DE, MG, BN, RG, and AR jointly wrote the manuscript. BN, EM, RD, DO, and HG reviewed the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, New York, USA and Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia, USA for providing the sorghum WGRS, RNASeq data, and phenotype data. The authors also acknowledge the supporting funds from AVISA (OPP1198373) and ICAR-BMGF (101165).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.666342/full#supplementary-material>

Supplementary Figure 1 | Whole genome sequence of 176 sorghum accessions mapped iteratively to the updated reference sequence assembly and the unmapped sequence reads were assembly iteratively. The plot represents the size of the sequence assembly gained from respective accessions.

Supplementary Figure 2 | The sorghum pangenome variable genes enrichment analysis and the metabolic pathways for (A) Biological process (B) Molecular function and (C) Cellular components. Top 10 GO terms identified for scoring GO terms for enrichment. Rectangles indicate the 10 most significant terms and the colour represents the relative significance ranging from red (most significant) to yellow (least significant). Each node has GO identifier and name with raw p-values and number of significant genes out of total genes annotated.

Supplementary Figure 3 | Insertions and deletions of various size distribution at sorghum pangenome **(A)** genome level and **(B)** gene level.

Supplementary Figure 4 | The genetic relationship of 354 sorghum accession neighbour joining (with 1000 bootstrap) analysis (the unrooted tree with branch length in red colour).

Supplementary Figure 5 | Sorghum 25 accessions (names and NCBI accessions) RNASeq read mapping density on pan-genome assembly for **(A–J)** Chromosome 1–10 **(K)** Scaffold sequence put together as single sequence and **(L)** Non reference sequence assembly contigs from sorghum accessions concatenated to single sequence as extra contig sequence.

Supplementary Table 1 | Raw sequence data used for sorghum pangenome and SNP calling analysis.

Supplementary Table 2 | The summary of gene presents in each sorghum accession.

Supplementary Table 3 | Gene count per each sorghum accession.

Supplementary Table 4 | Gene enrichment based on GO terms among the variable genes.

Supplementary Table 5 | Significantly enriched genes for BP, MF and CC components.

Supplementary Table 6 | Number of SNPs and density per each chromosome sequence.

Supplementary Table 7 | SNP count for pair-wise combination of sorghum accession.

Supplementary Table 8 | SNP count per sorghum accession.

Supplementary Table 9 | SNP validation with sorghum 3k SNP array.

Supplementary Table 10 | Sorghum genes unique to caudatum, durra, kafir, guniea.

Supplementary Table 11 | Genome wide association analysis of the SNPs showing the significant association with traits.

Supplementary Table 12 | Significantly associated SNPs from the genic region.

Supplementary Table 13 | Sorghum green leaf area trait data collection description.

Supplementary Table 14 | Drought RNASeq assay correspondence on sorghum pangenome extracontigs.

Supplementary Table 15 | Drought response genes count on pangenome assembly.

REFERENCES

- Abdel-Ghany, S. E., Ullah, F., Ben-Hur, A., and Reddy, A. S. N. (2020). Transcriptome analysis of drought-resistant and drought-sensitive sorghum (*Sorghum bicolor*) genotypes in response to PEG-induced drought stress. *Int. J. Mol. Sci.* 21:772. doi: 10.3390/ijms21030772
- Alexa, A., Rahnenführer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22, 1600–1607. doi: 10.1093/bioinformatics/btl140
- Andrews, S. (2015). FASTQC a quality control tool for high throughput sequence data. *Babraham Inst.* Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bayer, P. E., Golicz, A. A., Scheben, A., Batley, J., and Edwards, D. (2020). Plant pan-genomes are the new reference. *Nat. Plants* 6, 914–920. doi: 10.1038/s41477-020-0733-0
- Bayer, P. E., Hurgobin, B., Golicz, A. A., Chan, C. K. K., Yuan, Y., Lee, H. T., et al. (2017). Assembly and comparison of two closely related *Brassica napus* genomes. *Plant Biotechnol. J.* 15, 1602–1610. doi: 10.1111/pbi.12742
- Bekele, W. A., Wiekhorst, S., Friedt, W., and Snowdon, R. J. (2013). High-throughput genomics in sorghum: from whole-genome resequencing to a SNP screening array. *Plant Biotechnol. J.* 11, 1112–1125. doi: 10.1111/pbi.12106
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Brown, P. J., Myles, S., and Kresovich, S. (2011). Genetic support for phenotype-based racial classification in sorghum. *Crop Sci.* 51, 224–230. doi: 10.2135/cropsci2010.03.0179
- Bush, S. J., Castillo-Morales, A., Tovar-Corona, J. M., Chen, L., Kover, P. X., and Urrutia, A. O. (2014). Presence-absence variation in *A. thaliana* is primarily associated with genomic signatures consistent with relaxed selective constraints. *Mol. Biol. Evol.* 31, 59–69. doi: 10.1093/molbev/mst166
- Chen, W., Gao, Y., Xie, W., Gong, L., Lu, K., Wang, W., et al. (2014). Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nat. Genet.* 46, 714–721. doi: 10.1038/ng.3007
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly* 6, 80–92. doi: 10.4161/fly.19695
- Contreras-Moreira, B., Cantalapiedra, C. P., García-Pereira, M. J., Gordon, S. P., Vogel, J. P., Igartua, E., et al. (2017). Analysis of plant pan-genomes and transcriptomes with GET_HOMOLOGUES-EST, a clustering solution for sequences of the same species. *Front. Plant Sci.* 8:184. doi: 10.3389/fpls.2017.00184
- Cooper, E. A., Brenton, Z. W., Flinn, B. S., Jenkins, J., Shu, S., Flowers, D., et al. (2019). A new reference genome for *Sorghum bicolor* reveals high levels of sequence similarity between sweet and grain genotypes: implications for the genetics of sugar metabolism. *BMC Genomics* 20:420. doi: 10.1186/s12864-019-5734-x
- Cuevas, H. E., and Prom, L. K. (2020). Evaluation of genetic diversity, agronomic traits, and anthracnose resistance in the NPGS Sudan Sorghum Core collection. *BMC Genomics* 21:88. doi: 10.1186/s12864-020-6489-0
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330
- Dolatabadian, A., Bayer, P. E., Tirnaz, S., Hurgobin, B., Edwards, D., and Batley, J. (2020). Characterization of disease resistance genes in the *Brassica napus* pangenome reveals significant structural variation. *Plant Biotechnol. J.* 18, 969–982. doi: 10.1111/pbi.13262
- Evans, J., McCormick, R. F., Morishige, D., Olson, S. N., Weers, B., Hilley, J., et al. (2013). Extensive variation in the density and distribution of DNA polymorphism in sorghum genomes. *PLoS ONE* 8:e79192. doi: 10.1371/journal.pone.0079192
- Fracasso, A., Trindade, L. M., and Amaducci, S. (2016). Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE. *BMC Plant Biol.* 16:115. doi: 10.1186/s12870-016-0800-x
- Gao, L., Gonda, I., Sun, H., Ma, Q., Bao, K., Tieman, D. M., et al. (2019). The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat. Genet.* 51, 1044–1051. doi: 10.1038/s41588-019-0410-2
- Golicz, A. A., Batley, J., and Edwards, D. (2016b). Towards plant pangenomics. *Plant Biotechnol. J.* 14, 1099–1105. doi: 10.1111/pbi.12499
- Golicz, A. A., Bayer, P. E., Barker, G. C., Edger, P. P., Kim, H., Martinez, P. A., et al. (2016a). The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nat. Commun.* 7:13390. doi: 10.1038/ncomms13390
- Gordon, S. P., Contreras-Moreira, B., Woods, D. P., Des Marais, D. L., Burgess, D., Shu, S., et al. (2017). Extensive gene content variation in the *Brachypodium distachyon* pan-genome correlates with population structure. *Nat. Commun.* 8:2184. doi: 10.1038/s41467-017-02292-8
- Guo, H., Jiao, Y., Tan, X., Wang, X., Huang, X., Jin, H., et al. (2019). Gene duplication and genetic innovation in cereal genomes. *Genome Res.* 29, 261–269. doi: 10.1101/gr.237511.118
- Hart, G. E., Schertz, K. F., Peng, Y., and Syed, N. H. (2001). Genetic mapping of *Sorghum bicolor* (L.) Moench QTLs that control variation in tillering and other morphological characters. *Theor. Appl. Genet.* 103, 1232–1242. doi: 10.1007/s001220100582

- Hirsch, C. N., Foerster, J. M., Johnson, J. M., Sekhon, R. S., Muttoni, G., Vaillancourt, B., et al. (2014). Insights into the maize pan-genome and pan-transcriptome. *Plant Cell* 26, 121–135. doi: 10.1105/tpc.113.119982
- Hubisz, M. J., Falush, D., Stephens, M., and Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* 9, 1322–1332. doi: 10.1111/j.1755-0998.2009.02591.x
- Hurgobin, B., Golicz, A. A., Bayer, P. E., Chan, C. K. K., Tirnaz, S., Dolatabadian, A., et al. (2018). Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. *Plant Biotechnol. J.* 16, 1265–1274. doi: 10.1111/pbi.12867
- Jensen, S. E., Charles, J. R., Muleta, K., Bradbury, P. J., Casstevens, T., Deshpande, S. P., et al. (2020). A sorghum practical haplotype graph facilitates genome-wide imputation and cost-effective genomic prediction. *Plant Genome* 13:e20009. doi: 10.1002/tpg2.20009
- Johnson, S. M., Lim, F. L., Finkler, A., Fromm, H., Slabas, A. R., and Knight, M. R. (2014). Transcriptomic analysis of *Sorghum bicolor* responding to combined heat and drought stress. *BMC Genomics* 15:456. doi: 10.1186/1471-2164-15-456
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36. doi: 10.1186/gb-2013-14-4-r36
- Kimani, W., Zhang, L.-M., Wu, X.-Y., Hao, H.-Q., and Jing, H.-C. (2020). Genome-wide association study reveals that different pathways contribute to grain quality variation in sorghum (*Sorghum bicolor*). *BMC Genomics* 21:112. doi: 10.1186/s12864-020-6538-8
- Kong, L., Dong, J., and Hart, G. E. (2000). Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). *Theor. Appl. Genet.* 101, 438–448. doi: 10.1007/s001220051501
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., et al. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Sci. Adv.* 1:e1400218. doi: 10.1126/sciadv.1400218
- Letunic, I., and Bork, P. (2019). Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, R., Zhang, H., Zhou, X., Guan, Y., Yao, F., Song, G., et al. (2010). Genetic diversity in Chinese sorghum landraces revealed by chloroplast simple sequence repeats. *Genet. Resour. Crop Evol.* 57, 1–15. doi: 10.1007/s10722-009-9446-y
- Lin, K., Zhang, N., Severing, E. I., Nijveen, H., Cheng, F., Visser, R. G. F., et al. (2014). Beyond genomic variation - comparison and functional annotation of three *Brassica rapa* genomes: a turnip, a rapid cycling and a Chinese cabbage. *BMC Genomics* 15:250. doi: 10.1186/1471-2164-15-250
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444
- McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., et al. (2018). The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* 93, 338–354. doi: 10.1111/tj.13781
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., et al. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 28, 1297–1303. doi: 10.1101/gr.107524.110
- Montenegro, J. D., Golicz, A. A., Bayer, P. E., Hurgobin, B., Lee, H. T., Chan, C. K. K., et al. (2017). The pangenome of hexaploid bread wheat. *Plant J.* 90, 1007–1013. doi: 10.1111/tj.13515
- Morris, G. P., Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., et al. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 110, 453–458. doi: 10.1073/pnas.1215985110
- Motlaodi, T., Geleta, M., Bryngelsson, T., Fatih, M., Chite, S., and Ortiz, R. (2014). Genetic diversity in ex-situ conserved sorghum accessions of Botswana as estimated by: microsatellite markers. *Aust. J. Crop Sci.* 8, 35–43. Available online at: http://www.cropj.com/motlaodi_8_2014_35_43.pdf
- Obilana, A. B., Rao, E. P., Mangombi, N., and House, L. R. (1996). Classification of sorghum races in the Southern Africa Sorghum germplasm. *Sadq/Icrisat* 113–118. Available online at: <http://oar.icrisat.org/4740/>
- Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412
- Paterson, A. H., Bowers, J. E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457, 551–556. doi: 10.1038/nature07723
- Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174
- Price, H. J., Dillon, S. L., Hodnett, G., Rooney, W. L., Ross, L., and Johnston, J. S. (2005). Genome evolution in the genus *Sorghum* (Poaceae). *Ann. Bot.* 95, 219–227. doi: 10.1093/aob/mci015
- Ritter, K. B., McIntyre, C. L., Godwin, I. D., Jordan, D. R., and Chapman, S. C. (2007). An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* ssp. *bicolor* (L.) Moench, using AFLP markers. *Euphytica* 157, 161–176. doi: 10.1007/s10681-007-9408-4
- Sapkota, S., Boyles, R., Cooper, E., Brenton, Z., Myers, M., and Kresovich, S. (2020). Impact of sorghum racial structure and diversity on genomic prediction of grain yield components. *Crop Sci.* 60, 132–148. doi: 10.1002/csc2.20060
- Saxena, R. K., Edwards, D., and Varshney, R. K. (2014). Structural variations in plant genomes. *Brief. Funct. Genomics* 13, 296–307. doi: 10.1093/bfpg/elu016
- Schatz, M. C., Maron, L. G., Stein, J. C., Hernandez Wences, A., Gurtowski, J., Biggers, E., et al. (2014). Whole genome *de novo* assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of aus and indica. *Genome Biol.* 15:506. doi: 10.1186/PREACCEPT-2784872521277375
- Smit, A. F. A., Hubble, R., and Green, P. (2000). RepeatMasker. *Biotech Software Internet Rep.* 1, 36–39. doi: 10.1089/152791600319259
- Tian, T., You, Q., Zhang, L., Yi, X., Yan, H., Xu, W., et al. (2016). SorghumFDB: sorghum functional genomics database with multidimensional network analysis. *Database* 2016:baw099. doi: 10.1093/database/baw099
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Upadhyaya, H. D., Reddy, K. N., Vetriventhan, M., Gummam, M. K., Irshad Ahmed, M., Manyasa, E., et al. (2017). Geographical distribution, diversity and gap analysis of East African sorghum collection conserved at the ICRISAT genebank. *Aust. J. Crop Sci.* 11, 424–437. doi: 10.21475/ajcs.17.11.04.pne330
- Usha Kiranmayee, K. N. S., Hash, C. T., Sivasubramani, S., Ramu, P., Amindala, B. P., Rathore, A., et al. (2020). Fine-mapping of sorghum stay-green qtl on chromosome10 revealed genes associated with delayed senescence. *Genes* 11, 2–26. doi: 10.3390/genes11091026
- Valluru, R., Gazave, E. E., Fernandes, S. B., Ferguson, J. N., Lozano, R., Hirannaiah, P., et al. (2019). Deleterious mutation burden and its association with complex traits in sorghum (*Sorghum bicolor*). *Genetics* 211, 1075–1087. doi: 10.1534/genetics.118.301742
- Varoquaux, N., Cole, B., Gao, C., Pierroz, G., Baker, C. R., Patel, D., et al. (2019). Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses. *Proc. Natl. Acad. Sci. U.S.A.* 116, 27124–27132. doi: 10.1073/pnas.1907500116
- Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557, 43–49. doi: 10.1038/s41586-018-0063-9
- Woldesemayat, A. A., and Ntwasa, M. (2018). Pathways and network based analysis of candidate genes to reveal cross-talk and specificity in the sorghum (*Sorghum bicolor* (L.) Moench) responses to drought and it's co-occurring stresses. *Front. Genet.* 9:557. doi: 10.3389/fgene.2018.00557

- Yan, S., Wang, L., Zhao, L., Wang, H., and Wang, D. (2018). Evaluation of genetic variation among sorghum varieties from southwest China via genome resequencing. *Plant Genome* 11:170098. doi: 10.3835/plantgenome2017.11.0098
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P. C., Hu, L., et al. (2016). Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* 48, 927–934. doi: 10.1038/ng.3596
- Yao, W., Li, G., Zhao, H., Wang, G., Lian, X., and Xie, W. (2015). Exploring the rice dispensable genome using a metagenome-like assembly strategy. *Genome Biol.* 16:187. doi: 10.1186/s13059-015-0757-3
- Zhao, J., Bayer, P. E., Ruperao, P., Saxena, R. K., Khan, A. W., Golicz, A. A., et al. (2020). Trait associations in the pangenome of pigeon pea (*Cajanus cajan*). *Plant Biotechnol. J.* 18, 1946–1954. doi: 10.1111/pbi.13354
- Zheng, L. Y., Guo, X., Sen, H.e, B., Sun, L. J., and Peng, Y., Dong, S.S., et al. (2011). Genome-wide patterns of genetic variation in sweet and grain sorghum (*Sorghum bicolor*). *Genome Biol.* 12:R114. doi: 10.1186/gb-2011-12-11-r114
- Zhou, P., Silverstein, K. A. T., Ramaraj, T., Guhlin, J., Denny, R., Liu, J., et al. (2017). Exploring structural variation and gene family architecture with *de novo* assemblies of 15 Medicago genomes. *BMC Genomics* 18:261. doi: 10.1186/s12864-017-3654-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Ruperao, Thirunavukkarasu, Gandham, Selvanayagam, Govindaraj, Nebie, Manyasa, Gupta, Das, Odeny, Gandhi, Edwards, Deshpande and Rathore. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Genome-Wide Analysis of *MADS-Box* Genes in Foxtail Millet (*Setaria italica* L.) and Functional Assessment of the Role of *SiMADS51* in the Drought Stress Response

Wan Zhao^{1,2†}, Li-Li Zhang^{1†}, Zhao-Shi Xu², Liang Fu³, Hong-Xi Pang¹, You-Zhi Ma^{2*} and Dong-Hong Min^{1*}

OPEN ACCESS

Edited by:

Wolfram Weckwerth,
University of Vienna, Austria

Reviewed by:

Jose Manuel Alvarez,
University of Oviedo, Spain
Yusuf Khan,
Inland Norway University of Applied
Sciences, Norway

*Correspondence:

You-Zhi Ma
mayouzhi@caas.cn
Dong-Hong Min
mdh2493@126.com

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Crop and Product Physiology,
a section of the journal
Frontiers in Plant Science

Received: 29 January 2021

Accepted: 26 April 2021

Published: 28 June 2021

Citation:

Zhao W, Zhang L-L, Xu Z-S, Fu L,
Pang H-X, Ma Y-Z and Min D-H
(2021) Genome-Wide Analysis
of *MADS-Box* Genes in Foxtail Millet
(*Setaria italica* L.) and Functional
Assessment of the Role of *SiMADS51*
in the Drought Stress Response.
Front. Plant Sci. 12:659474.
doi: 10.3389/fpls.2021.659474

¹ College of Agronomy, Northwest A&F University/State Key Laboratory of Crop Stress Biology for Arid Areas, Yangling, China, ² Institute of Crop Science, Chinese Academy of Agricultural Sciences/National Key Facility for Crop Gene Resources and Genetic Improvement, Key Laboratory of Biology and Genetic Improvement of Triticeae Crops, Ministry of Agriculture, Beijing, China, ³ Xinxiang Academy of Agricultural Sciences of He'nan Province, Xinxiang, China

MADS-box transcription factors play vital roles in multiple biological processes in plants. At present, a comprehensive investigation into the genome-wide identification and classification of *MADS-box* genes in foxtail millet (*Setaria italica* L.) has not been reported. In this study, we identified 72 *MADS-box* genes in the foxtail millet genome and give an overview of the phylogeny, chromosomal location, gene structures, and potential functions of the proteins encoded by these genes. We also found that the expression of 10 MIKC-type *MADS-box* genes was induced by abiotic stresses (PEG-6000 and NaCl) and exogenous hormones (ABA and GA), which suggests that these genes may play important regulatory roles in response to different stresses. Further studies showed that transgenic *Arabidopsis* and rice (*Oryza sativa* L.) plants overexpressing *SiMADS51* had reduced drought stress tolerance as revealed by lower survival rates and poorer growth performance under drought stress conditions, which demonstrated that *SiMADS51* is a negative regulator of drought stress tolerance in plants. Moreover, expression of some stress-related genes were down-regulated in the *SiMADS51*-overexpressing plants. The results of our study provide an overall picture of the *MADS-box* gene family in foxtail millet and establish a foundation for further research on the mechanisms of action of *MADS-box* proteins with respect to abiotic stresses.

Keywords: *MADS-box*, phylogenetic analysis, expression profiling, abiotic stress responses, foxtail millet

Abbreviations: ABA, abscisic acid; ABRE, ABA-responsive elements; CDS, coding sequence length; FPKM, fragments per kilobase of transcript per million mapped reads; FW, fresh weight; GA, gibberellic acid; GARE, GA responsive elements; Ka, non-synonymous substitution rate; Ks, synonymous substitution rate; LTRE, low-temperature responsive element; MDA, malonaldehyde; NCBI, National Center for Biotechnology Information; MeJA, methyl jasmonate; POD, peroxidase; RT-PCR, reverse transcription-PCR; qPCR, quantitative real-time-PCR; RT-qPCR, quantitative real-time RT-PCR; SA, salicylic acid; SOD, superoxide dismutase; TF, transcription factor.

INTRODUCTION

Transcription factors play multiple roles over the entire life cycle of higher plants (Riechmann and Ratcliffe, 2000; Singh et al., 2002). In combination with *cis*-regulatory sequences, TFs activate or inhibit the expression of specific target genes in different tissues or cells, or in response to environmental conditions, and thus participate in the growth, development, morphogenesis, and biotic and abiotic stress responses process in plants (Liu et al., 1999). According to the PlantTFDB4.0¹, 320,370 TFs from 165 plant species have been classified into 58 families (Jin et al., 2018). Among these families, MADS-box proteins comprise a large TF family and are ubiquitous in the plant kingdom (Alvarez-Buylla et al., 2008; Zhang et al., 2018). Based on the evolutionary relationships and sequence characterization, Alvarez-Buylla et al. (2008) classified MADS-box proteins into two major types, Type I and Type II, and they all have one thing in common; both types contain a MADS-box domain. Normally, Type I *MADS-box* genes in plants contain one or two exons, none or one intron, and the proteins encoded usually contain a highly conserved SRF-like MADS domain but lack a K domain (De Bodt et al., 2003b; Gramzow and Theißen, 2010). However, Type II *MADS-box* genes contain multiple introns and exons, and the corresponding proteins possess four domains; from the N to the C terminus these are a conserved MEF2-like MADS (M) domain, a less-well-conserved intervening (I) domain, a semi-conserved keratin-like (K) domain, and a C-terminal (C) domain and are thus also known as MIKC-type MADS-box proteins (De Bodt et al., 2003a; Kaufmann et al., 2005). The M domain is the most conserved region that enables the functions of nuclear localization, dimerization, DNA binding, and accessory factor binding (McGonigle et al., 1996; Riechmann et al., 1996; Immink et al., 2002; Gramzow et al., 2010). The I domain contributes to MADS-box protein dimerization (Kaufmann et al., 2005), and the K domain plays an important role in the formation of higher-order complexes in addition to dimerization (Yang and Jack, 2004; Puranik et al., 2014; Theißen et al., 2016). The C domain is the most variable region and is responsible for transcriptional activation (Honma and Goto, 2001). On account of the diversity of the I and K domains, Type II MADS-box proteins are divided into two clades, MIKC* and MIKC^c, in which MIKC^c proteins have a shorter I domain and a more conserved K domain than MIKC* proteins (Henschel et al., 2002). Furthermore, based on phylogenetic analysis, the MIKC^c proteins from angiosperms group into at least 14 distinct subclades; AG/STK (AGL11), AGL6, AGL12, AGL17, Bsister (GGM13), FLC, AP1 (SQUA), AP3 (DEF), PI (GLO), OsMADS32-like, SVP (StMADS11), SOC1(TM3), TM8, and SEP (Becker and Theißen, 2003; Zhao et al., 2006; Gramzow and Theißen, 2015). Within the same subgroup, *MADS-box* genes mostly share similar expression patterns and the proteins perform highly related functions.

As transcriptional regulators, *MADS-box* genes play crucial roles in ontogeny and signal transduction in higher plants (Schilling et al., 2018). At present, there are few published studies

on Type I *MADS-box* genes, although the results of Masiero et al. (2011) indicate that Type I *MADS-box* genes are significant in plant reproductive development. Research has shown that the expression of Type I *MADS-box* genes in *Arabidopsis* can affect development of the female gametophyte, embryo, and endosperm (Day et al., 2008; Walia et al., 2009; Tiwari et al., 2010; Wuest et al., 2010). In contrast, studies on the plant-specific Type II *MADS-box* genes have been more thorough and extensive, and have shown that such genes play a broader regulatory role in plants. Firstly, the Type II *MADS-box* genes are closely related to flower development, and studies of the Type II MADS genes from several floral homeotic mutants in dicots have led to the establishment of the well-known ABCDE model for the determination of floral organs (Krizek and Fletcher, 2005). Additionally, multiple organs, including the roots, leaves, buds, embryos, and seeds were found to express Type II *MADS-box* genes, which provides further evidence for their diverse roles in plant development (Fornara et al., 2004; Gramzow and Theißen, 2010). For example, the expression of *MdDAM1*, a MIKC-type *MADS-box* gene from apple, was restricted to buds and could control growth cessation and bud dormancy (Moser et al., 2020). A recent study showed that under high temperature stress, AGAMOUS-LIKE67 (AGL67) could be combined with EARLY BOLTING IN SHORT DAY (EBS) to negatively control seed germination through the zinc-finger protein SOMNUS (SOM) in *Arabidopsis* (Li et al., 2020). With the increase in the number of research studies, a large number of MIKC-type genes have been identified from different species and shown to be involved in processes related to the stress response and hormone effects (Arora et al., 2007; Jia et al., 2018; Wang et al., 2018; Wei et al., 2018). For example, the MIKC-type MADS-box transcription factor SVP2 from kiwifruit (*Actinidia* sp.) was found to participate in ABA-mediated dehydration pathways by modulating expression of numerous target genes (Wu et al., 2018). Expression of the AGL2-like gene *ZMM7-L* from maize was affected by a variety of stresses, including cold, salt, drought, and exogenous ABA (Zhang Z. et al., 2012). In rice (*Oryza sativa* L.), the expression of *OsMADS61* (*OsMADS26*, an AGL12 ortholog; Lee et al., 2008a), was enhanced in response to osmotic stress induced by D-mannitol, and further studies confirmed that overexpression of *OsMADS61* (*OsMADS26*; Khong et al., 2015) in rice had a negative impact on pathogen resistance and drought tolerance. Chen et al. (2019) found that the expression of an SEP/AGL2 subfamily *MADS-box* gene in pepper (*Capsicum annuum* L.) was induced by abiotic stresses (cold, heat, salt, and osmotic stress) and hormones (ABA, SA, and MeJA); in addition, the transgenic *CaMADS*-expressing plants were found to be more tolerant to low temperature, high salt, and mannitol treatments compared with WT plants (Chen et al., 2019). Also, more MIKC-type *MADS-box* genes associated with stress resistance have been identified in other species including wheat (*Triticum aestivum* L.), sheepgrass (*Leymus chinensis*), *Brassica rapa*, and tomato (*Solanum lycopersicum*) (Gopal et al., 2015; Guo et al., 2016; Jia et al., 2018; Schilling et al., 2020). These studies on the role of MIKC-type *MADS-box* genes on stress resistance provide crucial resources for potential use in plant breeding and crop improvement (Boden and Østergaard, 2019). At present,

¹<http://planttfdb.cbi.pku.edu.cn/>

genome-wide analyses and functional characterization of MADS-box proteins have been conducted in many organisms, but there are few reports describing *MADS-box* genes in foxtail millet.

In this study, we identified a total of 72 *MADS-box* genes (29 Type I and 43 Type II) from the foxtail millet genome and we performed a comprehensive analysis of these genes. Ten MIKC-type *MADS-box* genes from different clades were shown to be induced by various abiotic stresses (drought and salinity) and exogenous application of hormones (ABA and GA). The *SiMADS51* genes, which belongs to the AGL12 subgroup, was isolated from cDNA of the foxtail millet cultivar 'Yugu 1.' The expression of *SiMADS51* was induced by PEG-6000, NaCl, ABA, and GA treatments. Further studies showed that overexpression of *SiMADS51* reduced the drought resistance of transgenic *Arabidopsis* and rice plants. These results will be helpful in understanding the evolutionary relationships, gene structures, and biological functions of the MADS-box TFs in foxtail millet and will establish a foundation for elucidating the drought resistance mechanism of *SiMADS51* gene.

MATERIALS AND METHODS

Sequence Acquisition and Identification of *MADS-Box* Genes in Foxtail Millet

Hidden Markov Model (HMM) were employed to identify the *MADS-box* genes from foxtail millet (*Setaria italica* L.) genome. The implementation details were carried out as described by He et al. with some revisions (He et al., 2019). *Setaria italica* protein sequences were downloaded from the Phytozome V12.1 website² to build a local protein database (Goodstein et al., 2012). The HMM profile of the SRF-TF domain (PF00319) was obtained from the Pfam database³. To acquire the foxtail millet *MADS-box* genes, the HMM profile were used to search against the local protein database by HMMER3 software with $E\text{-value} < e^{-5}$ (Eddy and Pearson, 2011). Subsequently, all candidate proteins were checked for containing MADS domain by submitting them as search queries to the National Center for Biotechnology Information Conserved Domain Database (NCBI-CDD, Lu et al., 2019) and SMART (Ivica and Peer, 2017). *MADS-box* protein sequences of *Setaria viridis* were obtained by the same method above. *MADS-box* protein sequences from *Arabidopsis*, rice (*Oryza sativa* L.), and potato (*Solanum tuberosum*) were obtained from published studies (Parenicová et al., 2003; Arora et al., 2007; Gao et al., 2018). Ultimately, we acquired 72 *MADS-box* genes in foxtail millet, 76 in *Setaria viridis*, 75 in rice, 109 in *Arabidopsis* and 156 in potato and named them based on their chromosomal location order (Supplementary Table 1).

Sequence Characteristics and Phylogenetic Analysis of *MADS-Box* Proteins in Foxtail Millet

Information regarding chromosomal distribution, predicted ORF length, number of amino acids, pfams, and introns of the foxtail

millet *MADS-box* genes were downloaded from Phytozome V12.1. Isoelectric points (pI) and molecular weights (MW) were computed using the ProtParam tool on the ExPASy Server⁴. Full details are given in Table 1.

The *MADS-box* protein sequences were aligned using ClustalX 2.0 with the default parameters (Larkin et al., 2007). A phylogenetic tree was constructed using the maximum likelihood method with 1,000 bootstrap replicates as implemented in MEGA 7 (Sudhir et al., 2016).

Chromosomal Localization and Gene Duplication Analysis of *MADS-Box* Genes in Foxtail Millet

All identified *MADS-box* genes were mapped to the nine foxtail millet chromosomes using the information obtained from the foxtail millet database using MapDraw software (Liu and Meng, 2003). Gene duplication analysis of foxtail millet *MADS-box* genes was conducted using the screening criteria proposed by Gu et al. (2002). The software KaKs_calculator (Zhang et al., 2006) was used to calculate the Ka/Ks values.

Predicted Protein Motifs and Gene Structure Characterization of Foxtail Millet *MADS-Box* Genes

The Multiple EM for Motif Elicitation (MEME) online tool⁵ was used to discover motifs in the predicted foxtail millet *MADS-box* proteins (Bailey et al., 2015). A MEME search was executed with the following parameters: motif count is 20; motif width ranges from 6 to 200 (inclusive) amino acids, and any number of repetitions (Arora et al., 2007). The detailed amino acid sequences of the 20 motifs are shown in Supplementary Table 2.

The coding sequences (CDS) and genomic DNA (gDNA) sequences of foxtail millet *MADS-box* genes were downloaded from the Phytozome database in order to predict gene structures. The Gene Structure Display Server 2.0 (GSDS 2.0⁶) online website was used to display the structures of the *MADS-box* genes (Hu et al., 2014).

Cis-Regulatory Element Analysis and Expression Analysis of Foxtail Millet *MADS-Box* Genes

In general, the 2,000 bp sequence located upstream of the initiation codon of a gene is considered to be the promoter region (Zhao et al., 2016). Therefore, the promoter sequences of all foxtail millet *MADS-box* genes were acquired from the Phytozome website. The primary *cis*-regulatory elements (CREs) in the promoter regions were then predicted using the online tool New PLACE⁷ (Higo et al., 1999). All predicted CREs obtained were counted and classified (Supplementary Table 3). The distribution of *cis*-elements related to abiotic stresses in the promoter regions were shown in Table 3.

⁴<https://web.expasy.org/protparam/>

⁵<http://meme-suite.org/tools/meme>

⁶<http://gsds.gao-lab.org/>

⁷<https://www.dna.affrc.go.jp/PLACE/?action=newplace>

²<https://phytozome.jgi.doe.gov/pz/portal.html/>

³<http://pfam.xfam.org/>

TABLE 1 | The detailed information of 72 *MADS-box* genes identified in foxtail millet genome.

Gene Name	Transcript name	Chr	Pfam	Nucleotide length (bp)	Amino acid length (aa)	pI	MW (kDa)	Intron number	Subfamily
<i>SiMADS01</i>	Seita.1G003500.1	1	PF00319; PF01486	1425	474	8.73	53.18	7	SOC1(TM3)
<i>SiMADS02</i>	Seita.1G072200.1	1	PF00319; PF01486	750	249	6.14	28.11	5	Bsisiter
<i>SiMADS03</i>	Seita.1G077600.1	1	PF00319	966	321	5.94	34.22	0	M
<i>SiMADS04</i>	Seita.1G148200.1	1	PF00319	441	146	4.76	16.05	1	M
<i>SiMADS05</i>	Seita.1G183300.1	1	PF00319	777	258	7.72	27.48	0	M
<i>SiMADS06</i>	Seita.1G209300.1	1	PF00319; PF01486	723	240	8.85	27.45	7	AGL17
<i>SiMADS07</i>	Seita.1G272300.1	1	PF00319	693	230	8.68	24.24	0	M
<i>SiMADS08</i>	Seita.1G273400.1	1	PF00319; PF01486	762	253	8.85	28.81	7	AGL6
<i>SiMADS09</i>	Seita.1G308200.1	1	PF00319; PF01486	729	242	9.27	27.64	7	AGL17
<i>SiMADS10</i>	Seita.1G328500.1	1	PF00319; PF01486	687	228	5.41	25.44	7	SVP
<i>SiMADS11</i>	Seita.2G002300.1	2	PF00319; PF01486	810	269	9.31	30.85	7	AP1(SQUA)
<i>SiMADS12</i>	Seita.2G026600.1	2	PF00319	1182	393	5.76	43.49	1	M
<i>SiMADS13</i>	Seita.2G086800.1	2	PF00319	741	246	5.64	26.81	0	M
<i>SiMADS14</i>	Seita.2G115700.1	2	PF00319	333	110	4.79	11.96	0	M
<i>SiMADS15</i>	Seita.2G266600.1	2	PF00319; PF01486	729	242	9.17	27.79	7	SEP
<i>SiMADS16</i>	Seita.2G383000.1	2	PF00319; PF01486	759	252	9.11	28.53	7	AP1(SQUA)
<i>SiMADS17</i>	Seita.3G055200.1	3	PF00319	1434	477	4.41	51.33	1	M
<i>SiMADS18</i>	Seita.3G073000.1	3	PF00319; PF01486	867	288	9.23	32.59	7	AG
<i>SiMADS19</i>	Seita.3G098400.1	3	PF00319; PF01486	795	264	9.1	29.85	6	AG
<i>SiMADS20</i>	Seita.3G098800.1	3	PF00319; PF01486	648	215	9.34	24.66	7	AGL12
<i>SiMADS21</i>	Seita.3G236800.1	3	PF00319; PF01486	648	215	8.45	25.11	6	PI(GLO)
<i>SiMADS22</i>	Seita.3G278000.1	3	PF00319	294	97	4.45	10.85	1	M
<i>SiMADS23</i>	Seita.3G280400.1	3	PF00319	588	195	10.25	21.22	0	M
<i>SiMADS24</i>	Seita.3G301600.1	3	PF00319	471	156	5.44	17.25	0	M
<i>SiMADS25</i>	Seita.3G358100.1	3	PF00319; PF01486	735	244	7.07	27.42	7	AP1(SQUA)
<i>SiMADS26</i>	Seita.4G062600.1	4	PF00319; PF01486	684	227	7.59	26.22	7	SEP
<i>SiMADS27</i>	Seita.4G077200.1	4	PF00319; PF01486	669	222	5.36	24.4	6	SVP
<i>SiMADS28</i>	Seita.4G093200.1	4	PF00319	1191	396	4.99	43.22	9	MIKC*
<i>SiMADS29</i>	Seita.4G160200.1	4	PF00319	699	232	10.27	25.94	2	M
<i>SiMADS30</i>	Seita.4G163500.1	4	PF00319	645	214	9.59	23.2	0	M
<i>SiMADS31</i>	Seita.4G177800.1	4	PF00319	552	183	7.91	20.08	0	M
<i>SiMADS32</i>	Seita.4G184600.1	4	PF00319	510	169	5.44	18.55	0	M
<i>SiMADS33</i>	Seita.4G219100.1	4	PF00319	468	155	5.11	17.11	1	M
<i>SiMADS34</i>	Seita.4G238000.1	4	PF00319	960	319	5.62	34.27	0	M
<i>SiMADS35</i>	Seita.4G268200.1	4	PF00319; PF01486	717	238	4.97	26.54	4	Bsisiter
<i>SiMADS36</i>	Seita.4G277600.1	4	PF00319; PF01486	690	229	8.74	26.11	6	AP3(DEF)
<i>SiMADS37</i>	Seita.5G033100.1	5	PF00319	753	250	9.61	27.8	0	M
<i>SiMADS38</i>	Seita.5G036500.1	5	PF00319; PF01486	690	229	7.73	25.99	6	AGL17
<i>SiMADS39</i>	Seita.5G101300.1	5	PF00319; PF01486	507	168	8.9	18.66	4	FLC
<i>SiMADS40</i>	Seita.5G114500.1	5	PF00319	750	249	8.59	27.84	0	M
<i>SiMADS41</i>	Seita.5G143100.1	5	PF00319; PF01486	774	257	9.07	29.07	7	AG
<i>SiMADS42</i>	Seita.5G160200.1	5	PF00319	1347	448	5.25	46.65	0	M
<i>SiMADS43</i>	Seita.5G220600.1	5	PF00319	552	183	5.03	20.8	1	M
<i>SiMADS44</i>	Seita.5G225300.1	5	PF00319	1260	419	5.89	46.71	0	M
<i>SiMADS45</i>	Seita.5G303200.1	5	PF00319; PF01486	591	196	6.55	22.68	4	OsMADS32-like
<i>SiMADS46</i>	Seita.5G404600.1	5	PF00319; PF01486	630	209	7.1	24.21	6	PI(GLO)
<i>SiMADS47</i>	Seita.5G406700.1	5	PF00319; PF01486	810	269	9.11	29.89	6	AG
<i>SiMADS48</i>	Seita.5G424200.1	5	PF00319	984	327	5.18	34.94	0	M
<i>SiMADS49</i>	Seita.5G425300.1	5	PF00319	1389	462	5.11	49.4	0	M

(Continued)

TABLE 1 | Continued

Gene Name	Transcript name	Chr	Pfam	Nucleotide length (bp)	Amino acid length (aa)	pI	MW (kDa)	Intron number	Subfamily
<i>SiMADS50</i>	Seita.5G432700.1	5	PF00319	840	279	10.58	29.46	4	FLC
<i>SiMADS51</i>	Seita.6G013400.1	6	PF00319; PF01486	678	225	6.61	25.39	6	AGL12
<i>SiMADS52</i>	Seita.6G156800.1	6	PF00319; PF01486	702	233	8.89	26.76	6	AGL17
<i>SiMADS53</i>	Seita.6G194800.1	6	PF00319	1035	344	4.58	37.43	11	MIKC*
<i>SiMADS54</i>	Seita.6G223400.1	6	PF00319; PF01486	738	245	8.75	28.42	7	SEP
<i>SiMADS55</i>	Seita.6G223600.1	6	PF00319	528	175	8.97	19.88	4	FLC
<i>SiMADS56</i>	Seita.6G223700.1	6	PF00319	372	123	10.13	14.22	1	FLC
<i>SiMADS57</i>	Seita.7G040500.1	7	PF00319	456	151	4.49	16.64	1	M
<i>SiMADS58</i>	Seita.7G109700.1	7	PF00319; PF01486	609	202	7.06	23.29	6	AGL12
<i>SiMADS59</i>	Seita.7G110000.1	7	PF00319	330	109	9.94	12.49	2	AGL12
<i>SiMADS60</i>	Seita.7G125400.1	7	PF00319; PF01486	717	238	7.72	27.18	7	AGL17
<i>SiMADS61</i>	Seita.7G210200.1	7	PF00319; PF01486	765	254	8.59	28.48	7	AGL6
<i>SiMADS62</i>	Seita.7G235900.1	7	PF00319; PF01486	738	245	9.27	28.01	5	Bsisiter
<i>SiMADS63</i>	Seita.8G182900.1	8	PF00319	627	208	6.21	22.05	0	M
<i>SiMADS64</i>	Seita.8G220800.1	8	PF00319	1110	369	6.2	40.94	10	MIKC*
<i>SiMADS65</i>	Seita.9G088700.1	9	PF00319; PF01486	741	246	6.45	27.88	7	SEP
<i>SiMADS66</i>	Seita.9G088900.1	9	PF00319; PF01486	750	249	9.33	28.47	7	AP1(SQUA)
<i>SiMADS67</i>	Seita.9G237300.1	9	PF00319	1098	365	5.03	40.09	4	M
<i>SiMADS68</i>	Seita.9G270800.1	9	PF00319	1230	409	5.18	43.79	1	M
<i>SiMADS69</i>	Seita.9G342700.1	9	PF00319; PF01486	678	225	9.6	25.62	6	SOC1(TM3)
<i>SiMADS70</i>	Seita.9G393900.1	9	PF00319	717	238	5.6	26.57	0	M
<i>SiMADS71</i>	Seita.9G513900.1	9	PF00319; PF01486	687	228	8.47	25.63	7	SVP
<i>SiMADS72</i>	Seita.9G561000.1	9	PF00319	186	61	10.65	6.97	0	SOC1(TM3)

The RNA-seq data for foxtail millet plants under drought stress was extracted from our previous research to investigate the expression of *SiMADS*-box genes in response to drought (Yu et al., 2018). The transcriptome sequence data for different tissues including the leaves, roots, stems, and tassel inflorescences were downloaded from Expression Atlas⁸ (Zhang G. Y. et al., 2012). Using the FPKM values (fragments per kilobase of transcript per million mapped reads) we obtained, a heat map of *SiMADS-box* gene expression in the different tissues was drawn with EvolView (Zhang H. K. et al., 2012).

Plant Materials, Growth Conditions, and Treatments

In this study, wild type *Arabidopsis* (ecotype Col-0) and rice (cv. 'Nipponbare') were used as recipients in the transformation experiments. Foxtail millet variety 'Yugu1' was used for the cDNA amplification, promoter sequence amplification, and expression analysis of the *SiMADS51* gene. The seeds of foxtail millet were sown in nutrition soil and grown at 28°C in a greenhouse at ~65% relative humidity and a 14 h/10 h (day/night) photoperiod. Three-week-old seedlings were treated with various abiotic stresses and exogenous hormones, including drought (15% PEG-6000 to simulate drought conditions), NaCl (100 mM), ABA (100 μM), and GA (100 μM). Leaves of foxtail millet seedlings were sampled at 0, 1, 3, 6, 12, and 24 h

after the treatments and frozen immediately in liquid nitrogen for RNA extraction.

RNA Isolation, RT-PCR, and qPCR Analysis

A Plant Total RNA Kit (ZP405, Beijing Zoman Biotechnology Co., Ltd.) was used for total RNA isolation following the kit instructions. For cDNA synthesis we used the EasyScript® One-Step gDNA Removal and cDNA Synthesis SuperMix kit (AE311-02, Transgen, Beijing). TransStart® Top Green qPCR SuperMix (+Dye II) kit (AQ132-21, Transgen, Beijing) was used for qPCR assays to determine the relative level of gene expression. The foxtail millet *actin* gene (GenBank ID: AF288226), *Arabidopsis actin2* gene (At3g18780), and rice *actin* gene (*LOC_Os03g50885.1*) were used as internal controls for normalization of gene expression in the three species. The relevant gene-specific primers used in this study are given in **Supplementary Table 6**. The experimental data was analyzed using the $2^{-\Delta\Delta Ct}$ method of Livak and Schmittgen (2001).

Subcellular Localization

To investigate the subcellular localization, the coding sequences excluding the termination codons of 10 *SiMADS-box* genes were cloned into the plant expression vector p16318h-GFP. The control plasmid and the recombinant plasmids were introduced into rice protoplasts separately as described previously (Zhang et al., 2011). The transfected protoplasts were incubated for 18 h

⁸<https://www.ebi.ac.uk/gxa/home>

in darkness at 22°C, after which they were examined using a confocal laser scanning microscope.

Generation of Transgenic *Arabidopsis* and Rice Plants Overexpressing the *SiMADS51* Gene

To obtain transgenic plants, the full-length coding region of *SiMADS51* was amplified from cDNA prepared from foxtail millet RNA (leaves sampled at 0 h). The coding sequence of *SiMADS51* excluding the termination codon was amplified and cloned into the plant expression vectors pCambia1302 and pCambia1305. To generate transgenic *Arabidopsis* plants, the fusion plasmid *pCambia1302:SiMADS51* was transfected into *Agrobacterium tumefaciens* strain GV3101 which was then used to transform *Arabidopsis* Col-0 plants by floral dipping method (Clough and Bent, 1998). 0.5X MS solid medium (2.5 g/L phytigel) containing 30 mg/L hygromycin was used to screen transgenic *Arabidopsis* lines (Zhao et al., 2017). Transgenic *Arabidopsis* seedlings were cultured at 22°C with ~65% relative humidity and a 16 h light/8 h dark photoperiod in a climate-controlled chamber. To obtain transgenic rice plants, the fusion plasmid *pCambia1305:SiMADS51* was transformed into *Agrobacterium tumefaciens* strain EHA105, and the recombinant strain was used to transfect rice embryonic calli according to the procedure detailed in Sallaud et al. (2003). Transgenic rice plants were selected by hygromycin (50 mg/L) and planted in soil substrate in a containment greenhouse under daily cycle of 28°C 14 h light/22°C 10 h dark (Khong et al., 2015). All constructs were sequence-verified before they were used for plant transformation. *SiMADS51*-positive transgenic plants were screened using PCR and cultured to the T₃ generation. The expression levels of *SiMADS51* in the T₃-generation transgenic lines were determined by RT-qPCR, and the lines with the highest expression levels were then used for the further evaluation of drought resistance.

Drought-Resistance Assessment of the Transgenic Plants

For the seed germination test, seeds of the wild type (WT; Col-0) and three *SiMADS51*-overexpressing (OE) *Arabidopsis* lines were sterilized and sown on 0.5X MS medium with or without 6% PEG-6000 or 9% PEG-6000. All the seeds were vernalized at 4°C for 3 days in the dark before being transferred to a climate-controlled chamber at 22°C with ~65% relative humidity and a 16 h light/8 h dark photoperiod. The numbers of germinated seeds were recorded every 12 h.

For the root growth test, 8-day-old WT and OE *Arabidopsis* seedlings grown on 0.5X MS medium were transferred to 0.5X MS medium with or without 9% PEG-6000. The transferred seedlings were cultured in the climate chamber for a week. The total root lengths were then measured using a root system scanner.

To evaluate drought-resistance in adult-stage transgenic *Arabidopsis* plants, 8-day-old WT and OE *Arabidopsis* seedlings grown on 0.5X MS medium were transferred to soil with

suitable moisture. Water was withheld from 2-week-old soil-grown *Arabidopsis* seedlings until they wilted (about 2 weeks). The plants were then re-watered and allowed to recover for 1 week in the growth chamber.

To assay drought stress in transgenic rice plants, hydroponics experiment under PEG-6000 simulated drought conditions were conducted in the growth chamber. Rice seeds were sterilized with 2.5% sodium hypochlorite (NaClO) for 30 min and germinated in tap water at 28-degree incubator for 2 days. Germinated seeds were cultured in 0.5X MS liquid medium for 2 weeks under the conditions described above. The seedlings were then transferred to 0.5X MS liquid medium containing 15% or 20% PEG-6000 and continued to grow until the seedlings were all wilted (about 1 week). Next, the seedlings were transferred to 0.5X MS liquid medium and allowed to recover for 1 week. All liquid media were changed every 2 days.

Physiological Measurements

For physiological measurements, the leaf samples from the control and transgenic plants were collected before and after stress treatments. The measurements of proline (Pro), chlorophyll, and MDA contents and the assays of SOD and POD activity were conducted using the corresponding test kits (Comin, China) following the manufacturer's protocols. A Varioskan LUX Multimode Microplate Reader (Thermo Fisher Scientific, United States) was used to measure the absorbance values. All measurements were repeated three times with three biological replicates.

Data Analysis

All experiments above were replicated three times independently. GraphPad Prism 5.0 software was used for statistical analyses. Statistical analysis of the data was performed with Student's *t*-test and ANOVA (a one-way analysis of variance). The significant or extremely significant differences ($p < 0.05$, $p < 0.01$) between two sets data are marked with single (*) or double (**) asterisks, respectively, in the figures.

RESULTS

Identification, Characterization, and Phylogenetic Analysis of Foxtail Millet *MADS-Box* Genes

A total of 72 candidate genes encoding MADS-box domain (SRF-TF) were identified in the foxtail millet (*Setaria italica* L.) genome and named as *SiMADS01*~*SiMADS72* based on their chromosomal locations. The predicted open reading frames (ORFs) of the *SiMADS-box* genes ranged from 186 to 1,434 bp, and the encoded amino acid sequences varied from 61 to 477 aa. The predicted molecular weights (MW) of the *SiMADS-box* proteins ranged from 6.79 to 53.18 kDa, with predicted isoelectric points (pI) that ranged from 4.41 to 10.65 (Table 1).

To investigate the phylogenetic relationships among MADS-box proteins in monocotyledons and dicotyledons, a phylogenetic tree was constructed with 488 MADS-box protein

sequences from *Arabidopsis* (109), potato (*Solanum tuberosum*) (156), rice (*Oryza sativa* L.) (75), foxtail millet (72), and *Setaria viridis* (76). The details of each gene are shown in **Supplementary Table 1**. We also constructed a second phylogenetic tree with only the 72 foxtail millet MADS-box proteins. Based on amino acid sequence similarities and the previous classification of MADS-box proteins from *Arabidopsis* and rice (Parenicová et al., 2003; Arora et al., 2007), the 72 MADS-box genes in the foxtail millet genome were divided into two major phylogenetic clades: 29 Type I (M-type) and 43 Type II (MIKC-type) (**Figures 1, 2A**). The 43 Type II MADS-box genes were then further divided into the 14 major subclades: SEP (4 proteins), AGL6 (2), AP1/SQUA (4), FLC (4), SOC1/TM3 (3), SVP/StMADS11 (3), PI/GLO (2), AP3/DEF (1), AG/STK (4), AGL17 (5), AGL12 (4), Bsister/GGM13 (3), OsMADS32-like (1), and MIKC* (3) (**Figures 1, 2A**). As can be seen from **Figure 1**, most of the type I MADS-box genes from each species clustered into one clade, showing a sister-group relationship. All of the foxtail millet Type II proteins clustered with their counterparts from *Arabidopsis*, potato, *Setaria viridis*, and rice with high bootstrap support apart from SiMADS45, which has no known ortholog in *Arabidopsis* and potato. In particular, the AGL17, AGL12, Bsister, and PI (GLO) subclades are significantly expanded in foxtail millet, rice, and *Setaria viridis*, compared with *Arabidopsis* and potato (**Figure 1**).

Chromosomal Locations and Gene Duplication Analysis of Foxtail Millet MADS-Box Genes

The physical positions of the MADS-box genes on the foxtail millet chromosomes were visualized with Mapdraw software. As shown in **Figure 2A**, the 72 MADS-box genes are distributed on nine foxtail millet chromosomes. Chromosome 8 contains the fewest MADS-box genes (~2.8%), while chromosomes 5, 4, and 1 contain the most (19.4%, 15.3%, and 13.9%, respectively), and account for nearly half of all the MADS-box genes (**Figure 2B**). Also, we found that the Type I and Type II MADS-box genes showed an uneven distribution on the foxtail millet chromosomes. There were no Type I genes identified on chromosome 6, whereas Type II genes are distributed across all nine chromosomes (**Figures 2A,B**). The numbers of Type I and Type II genes on chromosomes 2, 5, and 8 are equal, while there are more Type II genes than Type I genes on chromosomes 1, 3, 7, and 9 (**Figure 2B**). Also, some MADS-box genes from the same subclades tend to cluster together in one region of the chromosome. For example, SiMADS55 and SiMADS56, and SiMADS58 and SiMADS59, which belong to the FLC and AGL12 subclades, respectively, are tightly linked on chromosomes 6 (SiMADS55/56) and 7 (SiMADS58/59) (**Figure 2A**).

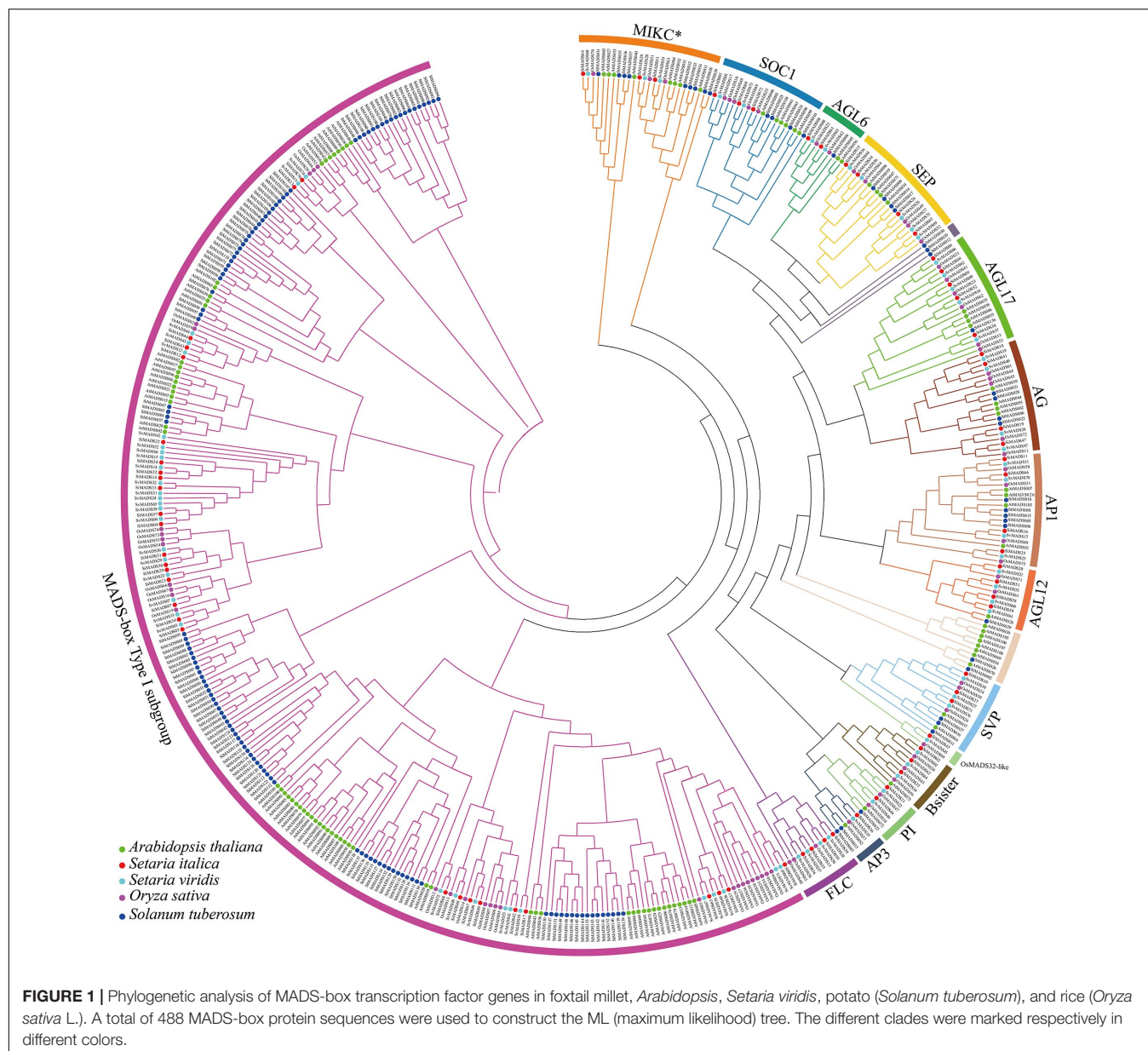
For gene duplication analysis, the coding sequences of the 72 foxtail millet MADS-box genes were used as queries in BLAST searches against all other SiMADS-box genes using an E -value $< 1e^{-10}$ and identity $> 85\%$. After the screening, only five pairs (SiMADS03 and SiMADS34, SiMADS06 and SiMADS60, SiMADS10 and SiMADS27, SiMADS21 and SiMADS46, and SiMADS33 and SiMADS57) met the search criteria, and both

members of each pair are located on different chromosomes (**Figure 2A** and **Table 2**). These results mean that the duplicated MADS-box genes are likely to have been generated by segmental duplication, which would produce many homologs on different chromosomes (Duan et al., 2015). The Ka, Ks, and Ka/Ks values of the homologs were calculated using a KaKs_calculator and are shown in **Table 2**. Based on the Ka/Ks ratios, we determined that the evolution of foxtail millet MADS-box genes was mainly accelerated by purifying selection (Hurst, 2002).

Conserved Protein Motifs and Structure of Foxtail Millet MADS-Box Protein Genes

The MEME on line tool (see text footnote 5) was used to identify the conserved motifs in the 72 predicted foxtail millet MADS-box proteins. A total of 20 conserved motifs (motif 1 to motif 20), were identified (**Figure 3B**). Detailed motif information is given in **Supplementary Table 2**. As can be seen in **Figure 3B**, Type I and Type II MADS-box proteins contain 15 and eight main motifs, respectively. The Type I MADS-box proteins exhibit more motif variation, and this is likely due to their non-conserved C-terminal regions. As anticipated, some particular motifs are specific to each family; for example, motifs 3, 9, 10, and 11 are found in Type I proteins, and motifs 2, 19, and 20 are found in Type II proteins (**Figures 3A,B**). The specificity of the protein motifs probably accounts for the functional diversity of the Type I and Type II family proteins. In general, the MADS-box proteins with similar motifs tended to cluster together in the phylogenetic analysis [the PI (GLO), AP1, and AGL6 subclades, for example], which implies that members of the same subclade have similar functions (Parenicová et al., 2003). However, we also found that some members, including SiMADS27, SiMADS35, and SiMADS59, differed from other members of the same subclade in terms of the particular motifs (**Figures 3A,B**). These differences might be due to the evolution of the MADS-box protein genes in foxtail millet.

To investigate the gene structures, the intron-exon organization of the foxtail millet MADS-box genes were analyzed using the GSDS online tool (see text footnote 6). As shown in **Figures 3A,C**, we found that the number of introns in foxtail millet MADS-box genes varied from 0 to 11, and the number of introns in Type I and Type II protein genes showed obvious differences. To be specific, 27 of 29 (93.1%) Type I genes contained either none or one intron, except for SiMADS29 and SiMADS67, while 40 of 43 (93%) Type II genes had at least four introns, although SiMADS56, SiMADS59, and SiMADS72 were among the exceptions (**Figures 3A,C**). This distribution of introns is analogous to that reported in *Arabidopsis* and rice (Parenicová et al., 2003; Arora et al., 2007). The PI (GLO), AP1 and AGL6 subclade proteins not only possess similar conserved motifs, but they also have similar numbers of introns, with the only differences between the members being the lengths of the introns and exons (**Figures 3A–C**). Also, within the same subclade, the proteins in which the motifs differ greatly from the others (e.g., SiMADS27, SiMADS35, and SiMADS59), the genes



also differ greatly in the arrangement of the introns and exons (Figures 3A–C).

Promoter *Cis*-Regulatory Element Enrichment Analysis of Foxtail Millet *MADS*-Box Genes

Regulation of gene expression is mainly dependent on the presence of CREs in the promoter region. To investigate the potential CREs present in the *MADS*-box gene family in foxtail millet, the 2000 bp of genomic DNA sequence upstream of the coding region of each gene was retrieved from the Phytozome website and searched against the New PLACE database. This analysis predicted 13,915 putative CREs in the 72 *SiMADS*-box genes (Supplementary Table 3). All the CREs can be divided

into four broad categories based on their functions and response to stimuli (Supplementary Table 3 and Figure 4). Statistical analysis showed that growth and biological process responsive elements comprised 48.95% of all CREs, followed by metabolism responsive elements (22.79%), and stress responsive elements (16.12%). Hormone responsive elements represented the lowest proportion (12.14%) of the CREs (Supplementary Table 3). For the growth and biological process responsive elements, light- and photosynthesis-related *cis*-elements accounted for 26.23% of the total (Figure 4A). Amongst the metabolism-related elements, the proportion of carbohydrate/sugar metabolism responsive elements (29.74%) followed by phenylpropanoid and flavonoid biosynthesis-related *cis*-elements (23.08%) was highest (Figure 4B). As shown in Figures 4C, numerous GA, Auxin, ABA, and SA response elements were identified upstream of

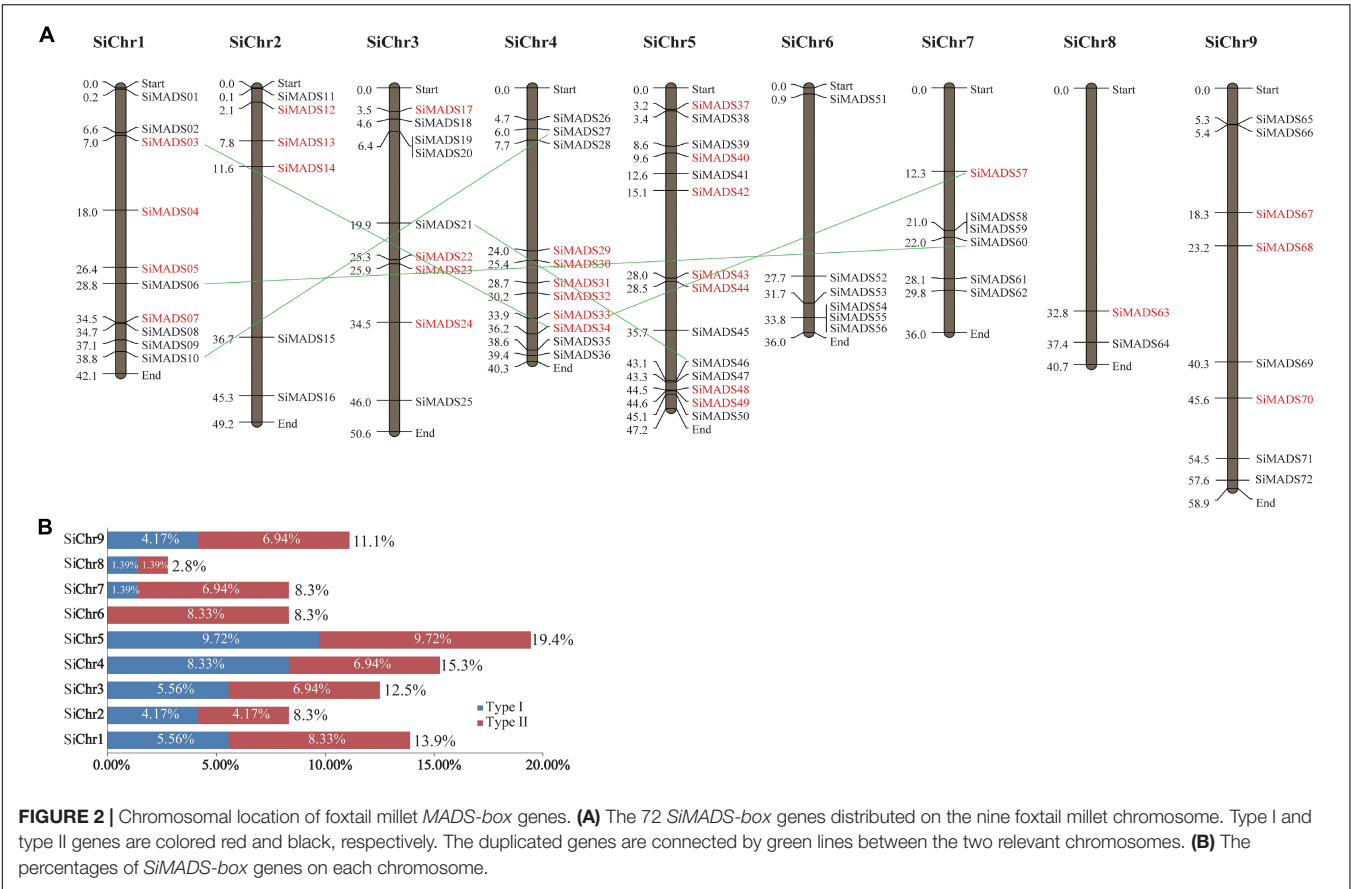


TABLE 2 | The *Ka*, *Ks*, and *Ka/Ks* values of the duplicated *SiMADS*-box genes.

Gene name	<i>Ka</i>	<i>Ks</i>	<i>Ka/Ks</i>	<i>P</i> -Value
<i>SiMADS03</i> and <i>SiMADS34</i>	1.13128	0.755894	1.49661	1.30E-12
<i>SiMADS06</i> and <i>SiMADS60</i>	0.130958	0.886409	0.14774	7.15E-41
<i>SiMADS10</i> and <i>SiMADS27</i>	0.810601	1.68021	0.48244	1.46E-24
<i>SiMADS21</i> and <i>SiMADS46</i>	0.0855293	1.82945	0.0467514	0
<i>SiMADS33</i> and <i>SiMADS57</i>	0.183993	0.336445	0.546873	0.00704695

the *SiMADS*-box genes, among which the GA response elements were the most abundant. Additionally, a large number of stress responsive *cis*-elements, which are involved in dehydration/water stress, wound signaling, and defense responsiveness, were also found in the *SiMADS*-box gene promoters (Figure 4D). These results imply the possible involvement of the *SiMADS*-box genes in plant development and stress tolerance.

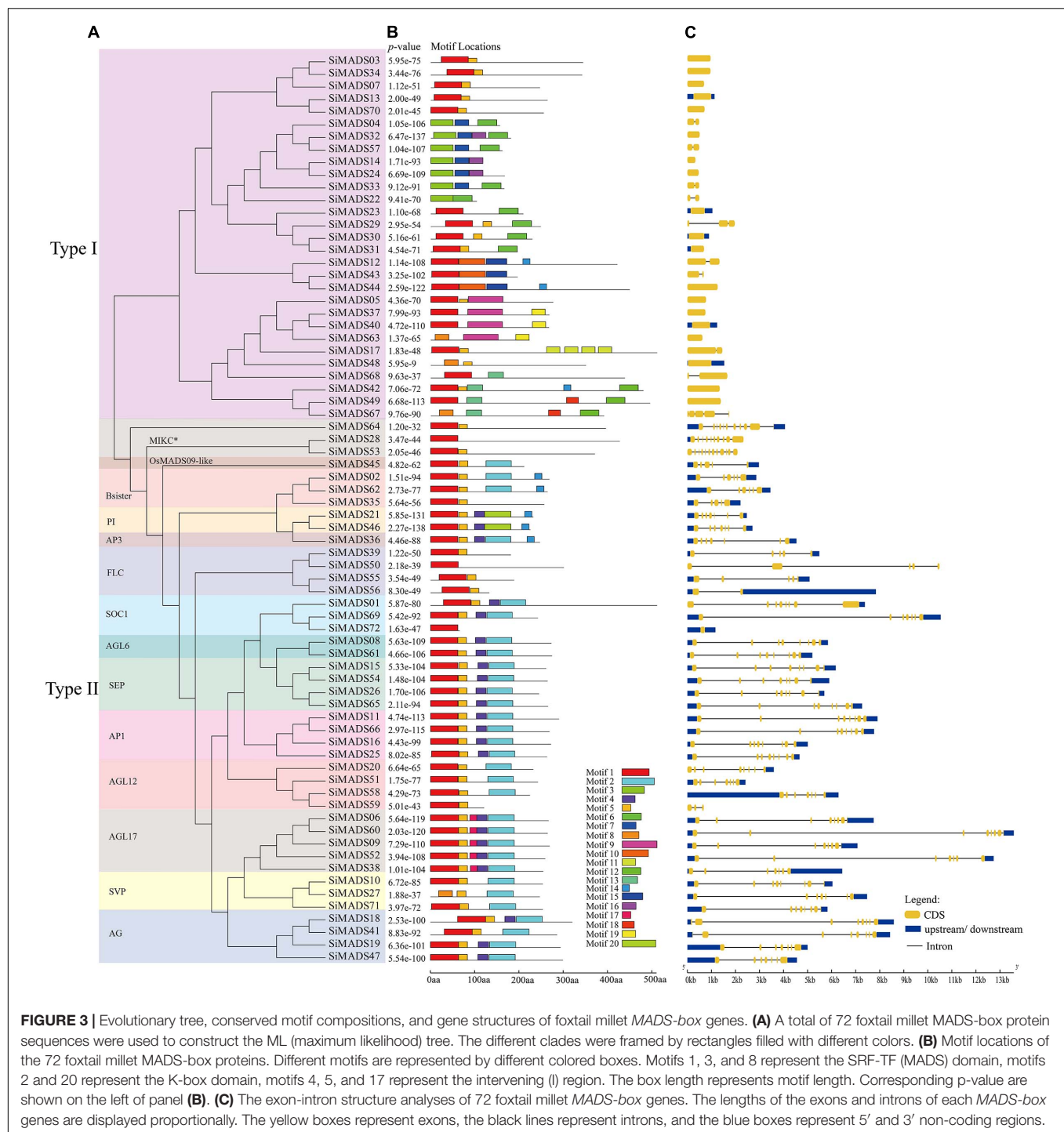
In our previous study, was performed RNA-seq analysis of drought-treated foxtail millet seedlings (Yu et al., 2018). Twenty-five *SiMADS*-box genes that showed a transcriptional response to drought stress were identified; among them, the relative expression levels of 10 *SiMADS*-box genes were found to be up-regulated in response to drought stress, and the differences were statistically significant ($P < 0.05$) (Supplementary Table 4). Interestingly, all of these 10 *SiMADS*-box genes belong to different Type II *MADS*-box subclades. Moreover, each of

the 10 *SiMADS*-box genes contained at least 21 stress- or hormone-related CREs in their promoter regions (Table 3). This demonstrates that a majority of Type II *MADS*-box genes in foxtail millet may be involved in the drought stress response.

Tissue-Specific Expression of *SiMADS*-Box Genes and Subcellular Localization of *SiMADS*-Box Proteins

To further analyze the expression patterns of the 10 Type II *SiMADS*-box genes in various tissues and organs, we downloaded the transcriptome sequencing data for different tissues of foxtail millet. The expression profiles of 10 *SiMADS*-box genes in leaf, root, stem and tassel inflorescence are given in Supplementary Table 5 and shown in a heatmap in Figure 5. The results show that the 10 *SiMADS*-box genes exhibited large differences in their expression profiles. For example, *SiMADS36*, *SiMADS51*, and *SiMADS55* showed their highest transcript levels in the tassel inflorescence, roots, and leaves, respectively, and were expressed at lower levels in other tissues; *SiMADS01* and *SiMADS10* were expressed at moderate to high levels in the leaves, roots, stems, and tassel inflorescence, while *SiMADS25* and *SiMADS28* were weakly expressed in all tissues and organs examined (Figure 5).

Subcellular localization information is important in the study of protein function. To investigate the subcellular distribution of the predicted *SiMADS*-box proteins, the coding sequences



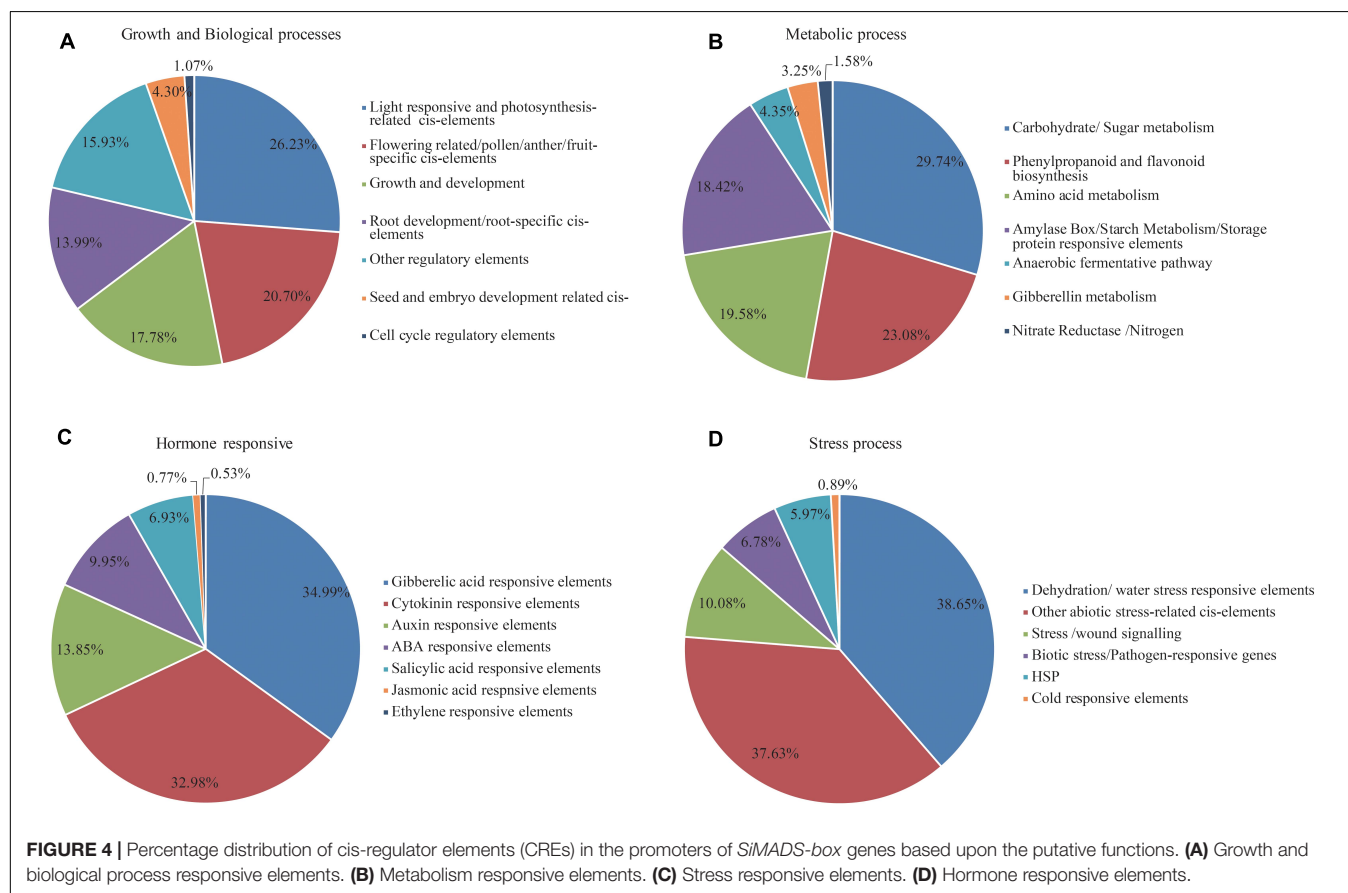
without the stop codons of the *SiMADS-box* genes were fused in frame with the gene for green fluorescent protein (GFP), and the constructs were transformed into rice protoplasts. We observed that the GFP fluorescence signals for the 10 *SiMADS-box* protein fusions were only found in the nucleus (**Supplementary Figure 1**), suggesting that these *SiMADS-box* proteins function mainly in the nucleus.

SiMADS-Box Gene Expression Is Induced by Multiple Abiotic Stresses

Previously published studies have suggested that *MADS-box* genes from different species are involved in the regulation of abiotic stresses and hormone response processes (Zhang G. Y. et al., 2012; Khong et al., 2015; Castelañ-Muñoz et al., 2019). To further investigate the possible involvement of the 10

TABLE 3 | Number of *cis*-elements in the promoter region of 10 *SiMADS*-box genes.

Element Name	<i>SiMADS01</i>	<i>SiMADS10</i>	<i>SiMADS25</i>	<i>SiMADS28</i>	<i>SiMADS36</i>	<i>SiMADS45</i>	<i>SiMADS51</i>	<i>SiMADS55</i>	<i>SiMADS56</i>	<i>SiMADS58</i>
ABREs	3	5	4	3	7	4	3	3	3	
MYB	4	3	3	6	9	7	6	3	2	14
MYC	10	2	18	7	7	10	4	9	9	17
W BOX	5	2	3	6	8	3	8	7	7	6
LTRE	1	2		3		1		2	1	1
BIHD1OS	3	1	1	4	2	1		3	1	3
CCAATBOX1	3	1	1	2		3	1	3	1	2
ACGTATERD1	5	1	4	2	5	7	4	2	7	
ELRECOREPCRP1	1						1		1	1
CBFHV	1	1	2	4		1		1		
GAREs	1				1	1				
GCCORE			3	2	1	2			2	
DRECRTCOREAT		1		2				1		
Total	37	21	39	41	40	40	27	34	34	44



SiMADS-box genes that are upregulated by drought in the response to abiotic stresses, we measured their expression patterns in response to PEG-6000, NaCl, ABA, and GA by RT-qPCR (Figure 6). The quantitative analysis results indicated that the 10 *SiMADS*-box genes had distinctly different transcriptional responses to the various abiotic stress or phytohormone treatments. Under the osmotic stress induced by PEG-6000, the transcription levels of *SiMADS51*, *SiMADS36*, and *SiMADS55*

were dramatically up-regulated (> 5-fold) and reached a peak at 12 h, 24 h, and 12 h, respectively (Figure 6A). In the salt stress treatment, the expression levels of *SiMADS01* and *SiMADS51* were up-regulated five-fold compared to the controls, and the expression of the two genes peaked at 6 h and 12 h, respectively (Figure 6B). Following ABA stress treatment, the expression of *SiMADS51*, *SiMADS56*, and *SiMADS10* was up-regulated four-fold, and the highest transcription levels occurred at 3 h, 6 h,

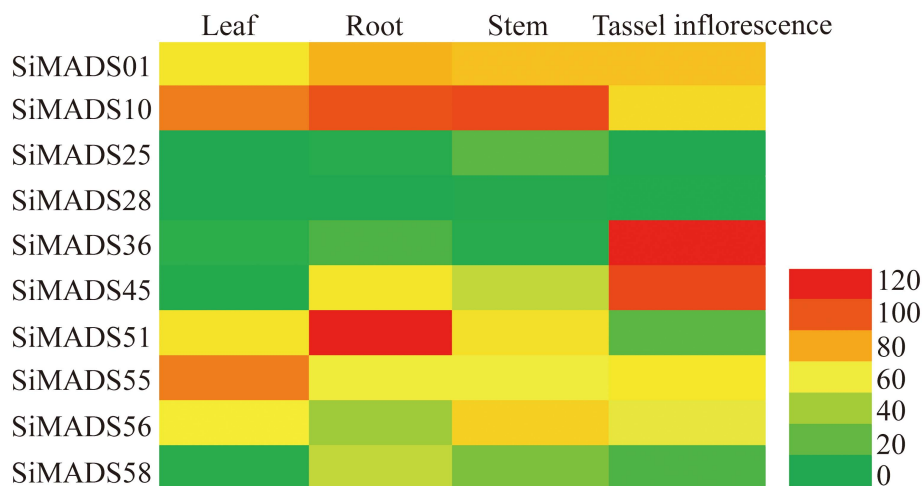


FIGURE 5 | The transcript accumulation of ten *SiMADS-box* genes in various foxtail millet tissues (root, leaf, stem, and tassel inflorescence). The transcriptome data of foxtail millet for different tissues were obtained from Expression Atlas. A heatmap was generated using EvolView. Transcript levels are indicated by different colors.

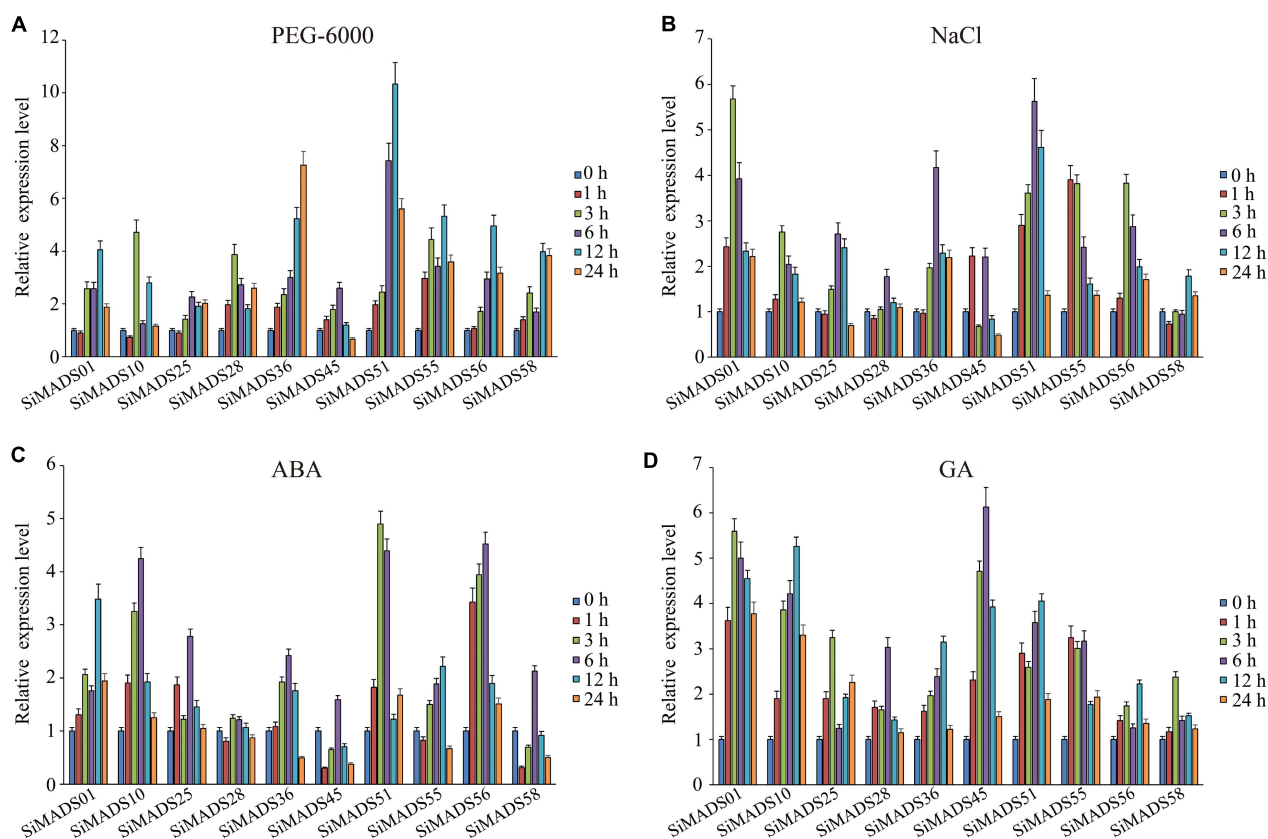


FIGURE 6 | Expression level of ten foxtail millet *MADS-box* genes under (A) PEG-6000, (B) NaCl, (C) ABA, and (D) GA treatment. The data were normalized to the foxtail millet *actin* gene (GenBank ID: AF288226). Three biological replicates were performed, and the values are presented as the means \pm SD.

and 6 h, respectively, while expression of the *SiMADS28* gene increased only slightly in response to ABA (<2-fold) (**Figure 6C**). When the seedlings were treated with GA, the transcript levels of

the 10 *SiMADS-box* genes were at least twice that of the control; *SiMADS45* showed the highest overall expression level, which occurred at 6 h (**Figure 6D**). These results suggest that foxtail

millet MADS-box TFs may play diverse roles in plant responses to abiotic stresses.

Overexpression of *SiMADS51* Affects Seed Germination in Response to PEG-6000 Treatment in *Arabidopsis*

Based on the gene expression analysis, *SiMADS51*, which belongs to the AGL12 subclade and is induced strongly by PEG-6000, was selected for further stress tolerance assays in transgenic *Arabidopsis* plants. The relative expression levels of *SiMADS51* gene in different line plants were shown in **Supplementary Figure 2A**. Homozygous T₃-generation transgenic *Arabidopsis* seeds were used in the germination test. After surface sterilization, the *Arabidopsis* seeds were sown on 0.5X MS medium with or without 6% PEG-6000 and 9% PEG-6000. Statistical analyses of the results showed that there were no significant differences in germination between wild-type (WT) and transgenic *Arabidopsis* lines (OE-1, OE-2, and OE-3) on 0.5X MS medium (**Figures 7A,B**), indicating that overexpression of *SiMADS51* probably has no effect on plant growth and development under normal conditions. In the presence of 6% PEG-6000 and 9% PEG-6000, germination of both wild-type and transgenic lines was inhibited; nonetheless, *SiMADS51*-overexpressing lines had lower germination percentages than WT (**Figures 7A,C,D**). For example, on 0.5X MS medium supplemented with 9% PEG-6000, the germination rates of seeds from each transgenic line within 2 days was 56.5% (OE-1), 51.9% (OE-2), and 55.6% (OE-3), significantly lower than the 73.1% germination rate in the wild-type seeds.

Overexpression of *SiMADS51* Reduced Drought Tolerance in Transgenic *Arabidopsis* Plants

To investigate the tolerance of *SiMADS51*-overexpressing lines to drought stress, transgenic *Arabidopsis* seedlings were cultured on 0.5X MS medium supplemented with PEG-6000 and also subjected to drought stress in soil. When grown on 0.5X MS medium, WT and *SiMADS51*-overexpressing (OE-1, OE-2, and OE-3) plants exhibited similar growth status (**Figure 8A**). However, on 0.5X MS medium supplemented with 9% PEG-6000, the roots of the transgenic line seedlings were shorter than those of the wild-type plants, and plants of the three transgenic lines had lower fresh weights compared to the wild-type plants (**Figures 8A–C**). When grown in soil, no significant differences in morphology were observed between wild-type plants and the three *SiMADS51*-overexpressing lines prior to drought treatment (**Figure 8D**). After drought stress treatment, seedlings of the *SiMADS51*-overexpressing lines appeared to be in poorer physical condition compared to the wild-type plants. As can be seen in **Figure 8D**, the OE line plants were badly wilted and bleached, while the WT plants were only slightly damaged. After re-watering, the survival rate of the wild-type *Arabidopsis* plants was higher than that of the three OE lines (**Figure 8E**). To study the possible physiological mechanism that explains the decreased drought tolerance in the *SiMADS51*-overexpressing lines, some stress-related physiological indicators were measured

for the OE lines and WT plants grown under normal and drought conditions. Under normal growth conditions, there were no significant differences in the proline and MDA contents between the OE and WT plants (**Figures 8F,G**). However, the proline content in the WT plants was higher than in the OE plants under drought stress conditions (**Figure 8F**). For MDA, the contents measured in detached leaves from the *SiMADS51* overexpressing lines were higher than in the WT plants under drought stress conditions (**Figure 8G**). These results indicate that decreased contents of stress-related metabolites and enhanced membrane lipid peroxidation may be the causes of reduced drought tolerance in the *SiMADS51*-overexpressing *Arabidopsis* plants.

Overexpression of *SiMADS51* Reduces Tolerance to Drought Stress in Transgenic Rice Plants

To gain further evidence of how *SiMADS51* overexpression reduces plant tolerance to drought stress, we transformed the *SiMADS51* gene into rice, and obtained 19 independent transgenic lines. The relative expression levels of *SiMADS51* gene in OE-5 and OE-12 line plants (T₃ generation) were higher than those of the other lines (**Supplementary Figure 2B**). Thus, the transgenic lines OE-5 and OE-12 were selected for further testing. PEG-6000 treatment was used to mimic dehydration stress conditions, and the drought tolerance of the two *SiMADS51*-overexpressing rice lines and control line (CK) plants were assessed at the vegetative growth stage. As shown in **Figure 9A**, all the seedlings exhibited similar growth status before stress treatment, while exposure to PEG-6000 stress for 1 week led to leaf wilting and even drying. However, in the 15% PEG-6000 stress treatment, leaves of the control plants showed delayed wilting and less curling compared to plants of the two transgenic OE lines (**Figure 9A**). After returning to normal growth conditions, only 57% and 49% of the transgenic OE-5 and OE-12 line seedlings lines recovered from the 20% PEG-6000 treatment, while 75% of the control seedlings survived (**Figure 9B**). In addition, the fresh weights of the transgenic line seedlings were less than those of the CK plants (**Figure 9C**). To investigate the physiological changes that occurred in the *SiMADS51*-OE and CK plants, the chlorophyll content and the POD and SOD activities were measured under drought stress and normal growth conditions. After PEG-6000 treatment, the POD and SOD activities increased significantly in both the *SiMADS51*-OE and CK plants compared with the activities in plants grown under normal conditions (**Figures 9D,E**). Under drought stress conditions, the POD and SOD activities in *SiMADS51*-OE plants were much lower than in the CK plants (**Figures 9D,E**). Furthermore, the chlorophyll contents of the *SiMADS51*-OE plants under PEG-6000 stress were also significantly lower than in the CK plants (**Figure 9F**). These results indicate that under drought stress, the photosynthesis rate and reactive oxygen species (ROS) scavenging capacity of the *SiMADS51*-OE plants were reduced compared to the CK plants, which led to the observed reduction in drought tolerance of the transgenic line plants.

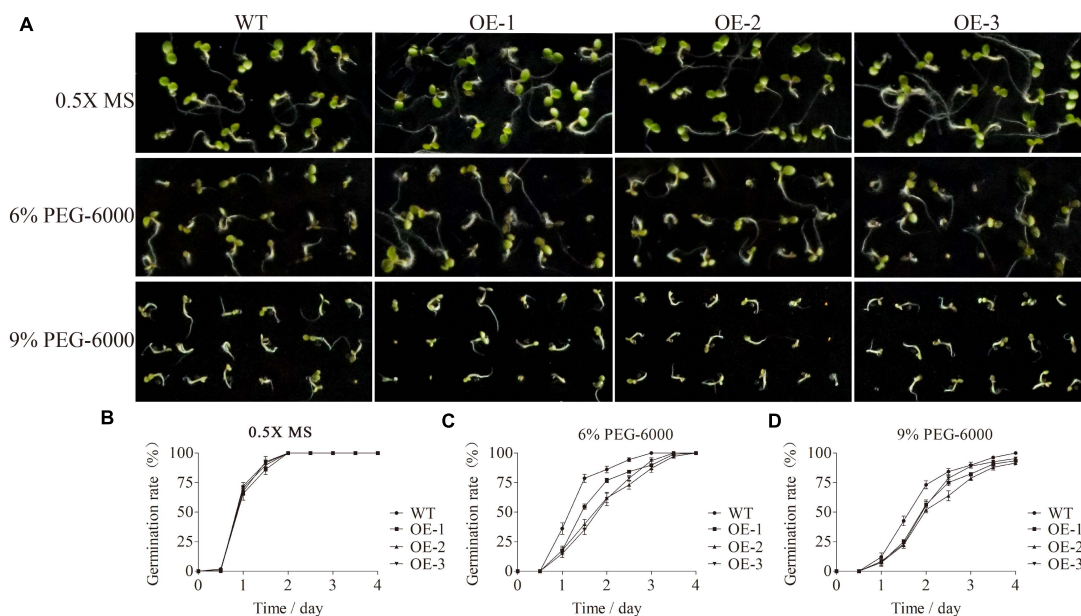


FIGURE 7 | Germination assays of wild-type (WT) and *SiMADS51* transgenic *Arabidopsis* seeds under PEG-6000 treatments. **(A)** Phenotypes of WT and *SiMADS51* transgenic *Arabidopsis* seeds on 0.5X Murashige and Skoog (MS) medium with or without 6% PEG-6000 and 9% PEG-6000, respectively. **(B–D)** Germination rates of WT and *SiMADS51* transgenic *Arabidopsis* seeds on 0.5X MS medium, 0.5X MS medium containing 6% PEG-6000, and 0.5X MS medium containing 9% PEG-6000 at different time points. Date for each time point are means of three biological replicates.

SiMADS51 Expression Alters the Transcription of a Group of Abiotic Stress-Related Genes in Rice

To explore the possible molecular mechanisms responsible for the reduced drought tolerance in transgenic plants, we examined the expression profiles of several rice genes reported to be related to abiotic stress tolerance. Leaves from normally grown plants were sampled for RT-qPCR analysis. As shown in **Figure 10**, compared with the control plants, the expression levels of *OsDREB2A* (Dubouzet et al., 2003), *OsMYB2* (Yang et al., 2012), and *OsPP2C06* and *OsPP2C49* (Singh et al., 2010), some signal transduction- and regulation-related genes known to be associated with abiotic stress tolerance, were down-regulated in the transgenic lines. Likewise, the expression level of two late stress-responsive genes, *OsLEA3* (Xiao et al., 2007) and *OsP5CS1* (Xiong et al., 2014), were also down-regulated in the transgenic lines. We also found that the expression of *OsNCED1* and *OsNCED3* (Hwang et al., 2010; Changan et al., 2018), two stress-related genes involved in ABA biosynthesis, was down-regulated in the transgenic plants grown under normal condition. These results indicate that *SiMADS51* may act as a negative regulator to inhibit the expression of abiotic stress-related genes, thereby impairing drought tolerance in plants.

DISCUSSION

Foxtail millet (*Setaria italica* L.) is one of the oldest cultivated crops in China (Jia et al., 2013). With the development of high

quality varieties and the improvement of cultivation techniques in recent years, improving the quality and production efficiency of millet will become a major research focus. Previous studies have shown that MADS-box TFs play important roles in plant growth, development, and the abiotic stress response (Honma and Goto, 2001; Becker and Theissen, 2003; Lee et al., 2008b; Khong et al., 2015). Therefore, the *MADS-box* gene family could be a key resource for promoting transgenic crops and traditional breeding to increase yield.

In this paper, a total of 72 *MADS-box* genes were identified in the foxtail millet genome (**Table 1**). In comparison with previously published studies, the number of *MADS-box* genes vary between monocotyledons and dicotyledons; for example, rice (*Oryza sativa* L.) (75), maize (*Zea mays* L.) (75), and *Sorghum bicolor* (65) (Arora et al., 2007; Zhao et al., 2011), all have fewer *MADS-box* genes than *Arabidopsis* (109), potato (*Solanum tuberosum*) (156), and tomato (*Solanum lycopersicum*) (131) (Gao et al., 2018; Wang et al., 2019). This suggests that monocotyledonous plants might have lost *MADS-box* genes during evolution. To examine the gain and loss of *MADS-box* genes in foxtail millet, a phylogenetic tree was constructed using the amino acid sequences of MADS-box proteins from foxtail millet, *Setaria viridis*, *Arabidopsis*, potato, and rice. As can be seen in **Figure 1** and **Supplementary Table 1**, the number of Type I *MADS-box* genes in *Arabidopsis* (63) and potato (116) was much higher than that in foxtail millet (43), rice (45), and *Setaria viridis* (43), while, the number of Type II *MADS-box* genes was similar in the five species. Therefore, we speculated that during the evolution of monocotyledons, the loss of *MADS-box* genes occurred mainly in the Type I clade. In foxtail millet, the Type

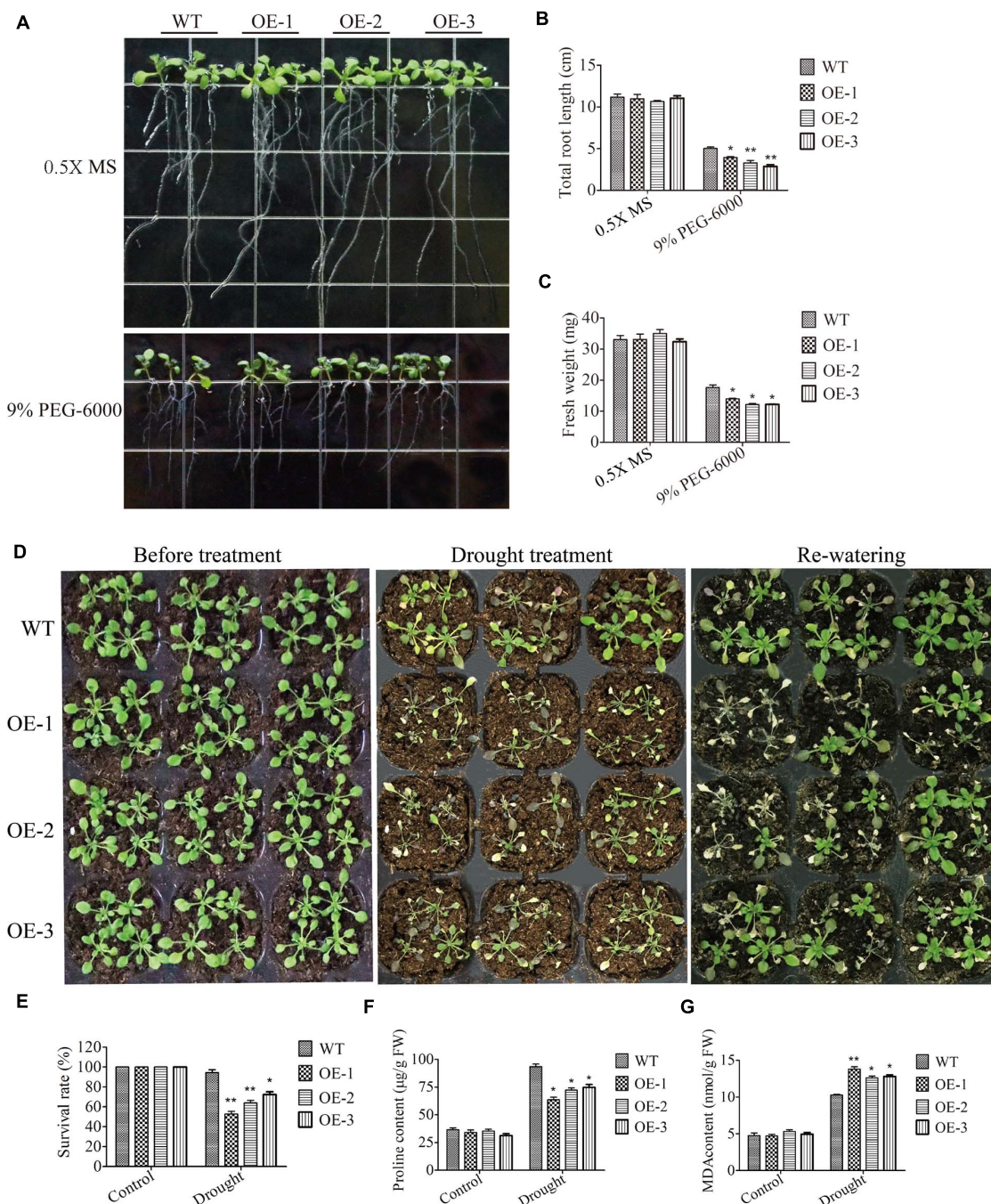


FIGURE 8 | Overexpression of *SiMADS51* reduced drought tolerance in transgenic *Arabidopsis* plants. **(A)** Root length assays of wild-type (WT) and *SiMADS51*-overexpressing plants under normal conditions and PEG-6000 treatment. **(B)** Total root lengths of seedlings. **(C)** Fresh weight of normal and PEG-6000 stressed plants. **(D)** Drought tolerance phenotypes of WT and *SiMADS51* transgenic *Arabidopsis* in soil. Three-week-old seedlings of WT and *SiMADS51*-overexpressing lines were dehydrated for 1 week and then rehydrated for 3 days. **(E)** Survival rate of normal and drought-stressed plants. **(F)** Proline and **(G)** MDA content were detected in WT and OE plants under drought or normal growth condition. Data are presented as the means \pm SDs of three independent replicates. The asterisks indicate significant differences between WT and *SiMADS51*-overexpressing lines (* $P < 0.05$; ** $P < 0.01$; Student's t -test).

II *MADS-box* genes could be divided into MIKC^c and MIKC^{*} clades, and the MIKC^c clade is further divided into 12 smaller subclades. In the OsMADS32-like subclade, the *MADS-box* genes are present in related monocot species but not in *Arabidopsis*

and potato, which suggests that these may have evolved after the divergence of monocots and dicots. In the AGL17, AGL12, Bsister, and PI (GLO) subclades, there are more *MADS-box* genes from rice, *Setaria viridis*, and foxtail millet than that from

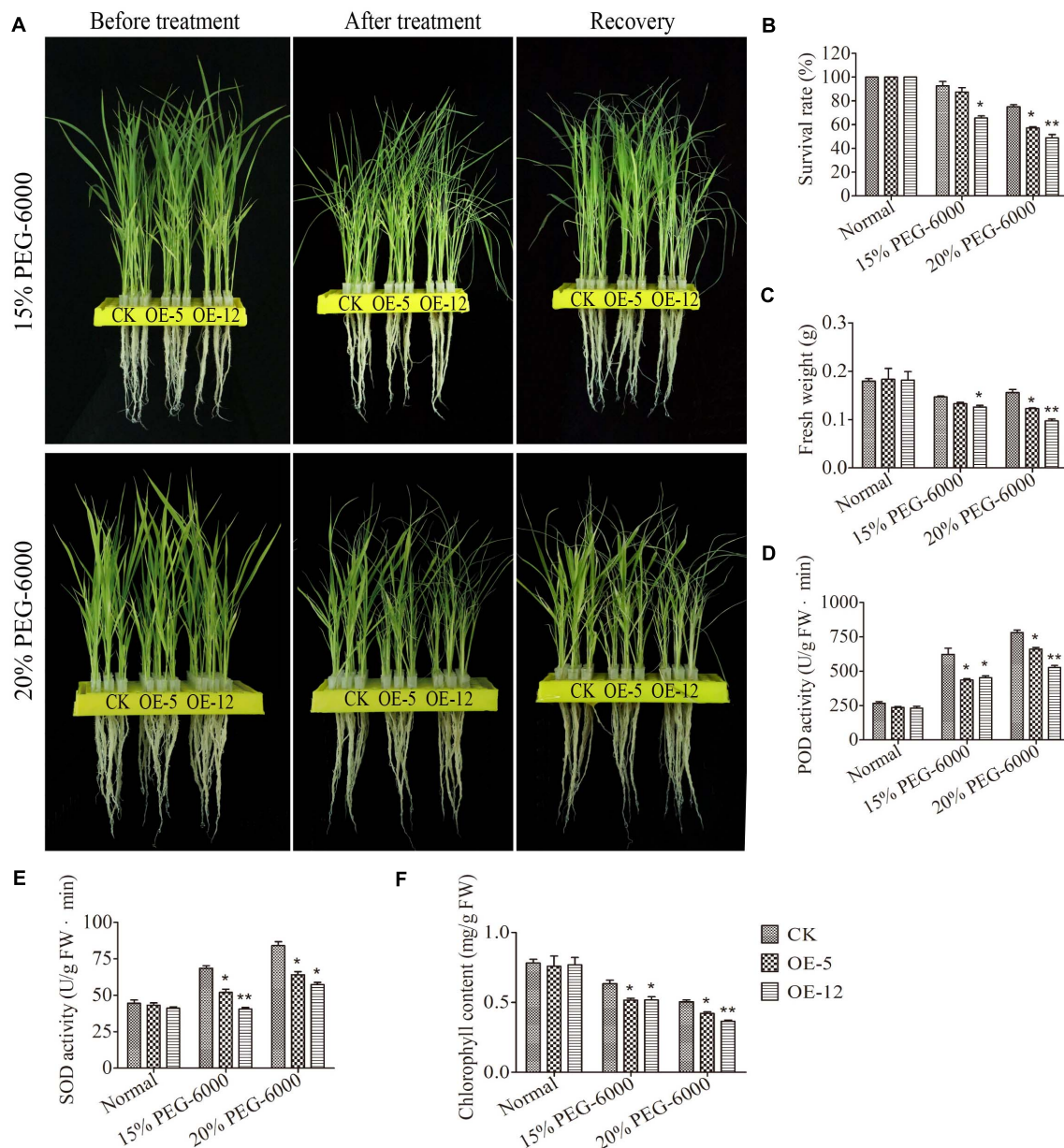


FIGURE 9 | Overexpression of *SiMADS51* reduces tolerance to drought stress in transgenic rice plants. **(A)** Phenotypes of CK and *SiMADS51*-transgenic rice under the treatment of PEG-6000. **(B)** Survival rate and **(C)** fresh weight of normal and PEG-6000 stressed plants. **(D)** POD, and **(E)** SOD activities were measured under PEG-6000 stress and normal growth conditions. **(F)** Chlorophyll content were detected in CK and OE plants under PEG-6000 stress or normal growth condition. Data are presented as the means \pm SDs of three independent replicates. The asterisks indicate significant differences between WT and *SiMADS51*-overexpressing lines (* $P < 0.05$; ** $P < 0.01$; Student's *t*-test).

Arabidopsis and potato, suggesting these subclades expanded in during the evolution of monocots.

The plant *MADS-box* gene family appears to have expanded mainly through tandem duplication events (Airoidi and Davies, 2012). It was previously reported that different groups of *MADS-box* genes underwent expansion in different plant species, such as M-type genes in rice, *SVP*-like genes in cotton, and *AGL17*-like and *Bsister*-like genes in wheat (*Triticum aestivum* L.) (Arora et al., 2007; Nardeli et al., 2018; Schilling et al., 2020). Our

analysis revealed that the amplification of foxtail millet *MADS-box* genes was due to segmental duplication alone (Figure 2 and Table 2). Also, the Ka/Ks ratios indicate that the *SiMADS-box* genes mainly experienced purifying selection (Ka/Ks ratio < 1) during evolution (Zhang et al., 2006). This evidence indicates that the expansion (and contraction) of the foxtail millet *MADS-box* gene family might be to allow it to adapt to changes in the environment, thereby contributing to its wider distribution (Theißen et al., 2018). In some cases, neofunctionalization and

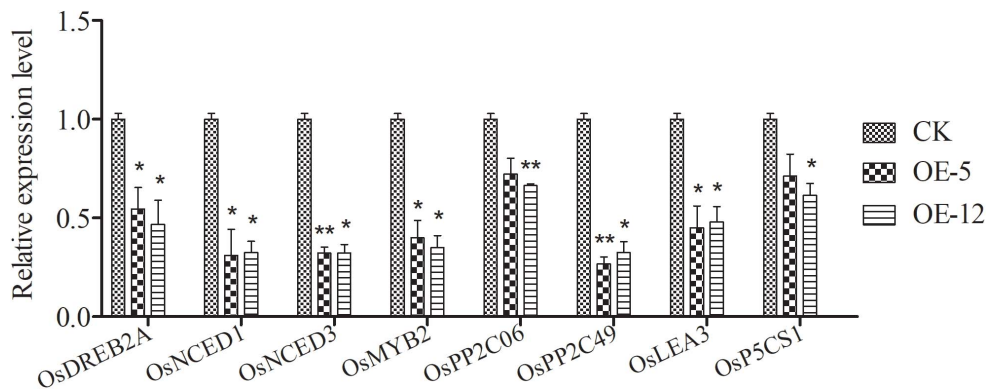


FIGURE 10 | Expression of several stress-related genes in *SiMADS51*-overexpressing plants (OE-5 and OE-12) and control line (CK) under normal condition. The expression of each gene was normalized using the *actin* gene (*LOC_Os03g50885*) as control. Data are presented as the means \pm SDs of three independent replicates. Asterisks indicate statistical significance (* $P < 0.05$; and ** $P < 0.01$; Student's *t*-test) compared with the corresponding controls.

subfunctionalization might be involved in gene duplication (Irish and Amy, 2005).

It is known that the structural diversity of genes drives the evolution of multigene families. Studies have shown that intron loss and insertion mutations are common during the evolution of plant *MADS-box* genes (Wei et al., 2015; Gao et al., 2018). In this study, we found that the number of introns in the *SiMADS-box* genes varied greatly (ranging from 0 to 11), most of the M-type *MADS-box* genes in foxtail millet had no introns, with a few exceptions, while the MIKC* genes had 9 to 11 introns, the most in any group of *SiMADS-box* genes, and was more than in other genes in the MIKC* subgroup (Table 1 and Figure 2C). Also, the protein motifs present in proteins the same family were not identical (Figure 2B). These differences among *MADS-box* genes in the same subgroup suggest that the loss or acquisition of introns may be a pattern of *SiMADS-box* gene evolution, and could be a major contributing factor to the functional diversity of the foxtail millet *MADS-box* family. Despite the differences, most of the *SiMADS-box* genes in the same group had similar intron-exon arrangements and the predicted proteins had a similar complement of motifs (Figures 2A–C). Similar results have been reported in monocots species such as rice (Arora et al., 2007) and wheat (Schilling et al., 2020), which imply evolutionary conservation within the *MADS-box* family.

Cis-regulatory elements play an important role in the regulation of plant gene expression; therefore, we conducted a systematic analysis of the CREs present in the promoter regions of the *SiMADS-box* genes. A large number of GA, ABA, defense, and dehydration responsive elements were identified in the *SiMADS-box* genes promoter region, which suggests that *SiMADS-box* genes might be regulated by various phytohormones, abiotic and biotic stresses. In addition, combined with our transcriptome sequencing data, we identified 10 foxtail millet *MADS-box* genes belonging to eight of MIKC-type subgroups in which expression was induced by drought stress. The promoter regions of these 10 *SiMADS-box* genes contain several stress-related *cis*-elements such as ABREs, GAREs, MYB, MYC, and W box (Table 3), indicating that

the expression of *SiMADS-box* genes can respond to multiple environmental cues. This is supported by further RT-qPCR analysis showing that the 10 *SiMADS-box* genes were induced significantly by PEG-6000, NaCl, ABA, and GA treatments (Figure 6). Moreover, tissue-specific expression patterns for the 10 *SiMADS-box* genes varied considerably in four plant organs (Figure 5), indicating the functional diversity of the *SiMADS-box* proteins. This knowledge could establish a foundation for further functional characterization of *MADS-box* gene family members in foxtail millet and other important crops.

In the present study, we observed that transgenic *Arabidopsis* and rice plants overexpressing *SiMADS51* showed reduced drought stress tolerance as indicated by the lower survival rates and poorer growth characteristics under drought stress conditions (Figures 7–9), which demonstrated that *SiMADS51* is a negative regulator of drought stress tolerance in plants. Previous studies have shown that *AtMADS028* (*XAL1/AGL12*; Tapia-López et al., 2008), an *Arabidopsis* gene homologous to *SiMADS51*, plays a role in regulating the root meristem, cell proliferation, and the flowering transition. The *SiMADS51*-homologous gene in rice, *OsMADS61* (*OsMADS26*; Khong et al., 2015), was found to negatively regulate resistance to pathogens and drought tolerance and had no significant impact on plant development. Homologous genes play different roles in monocotyledons and dicotyledons, which further illustrates the functional diversity of *MADS-box* genes. The mechanisms that underly the impaired drought tolerance in the *SiMADS51*-transgenic plants can be partially explained by the measured changes in several physiological and biochemical parameters. Generally, free proline is thought to be a compatible osmolyte that facilitates osmo-regulation, and it can help plants tolerate osmotic stress (Kochhar and Kochhar, 2005). The activities of POD and SOD can play a key role in protecting cell membranes against free radical attack and minimizing oxidative damage (Shah and Nahakpam, 2012). The chlorophyll content and MDA level can also be used as indicators of the degree of cellular injury (Gunes et al., 2011; Hong et al., 2016). In our study, we found that transgenic plants overexpressing *SiMADS51* had

higher MDA contents than did the control plants, while the free proline levels and chlorophyll contents and the POD and SOD activities in the *SiMADS51*-overexpressing plants were lower than those of the control plants under drought treatment. All of these results suggest that overexpression of *SiMADS51* reduced the tolerance of seedlings to drought stress, resulting in increased membrane permeability, reduced activity of peroxidases, and decreased photosynthesis. In addition, we also noticed that overexpression of *SiMADS51* in transgenic rice seedlings resulted in a decrease in the expression of a group of stress-responsive genes (**Figure 10**), including *OsPP2C06*, *OsPP2C49* (Singh et al., 2010), *OsDREB2a* (Dubouzet et al., 2003), *OsMYB2* (Yang et al., 2012), *OsLEA3* (Xiao et al., 2007), *OsP5CS1* (Xiong et al., 2014), *OsNCED1* (Changan et al., 2018), and *OsNCED3* (Hwang et al., 2010), which have been previously reported to be involved in the stress response. Hence, we can speculate that the decrease in plant stress tolerance may be due to *SiMADS51* directly or indirectly inhibiting the expression of these genes. Nevertheless, it is not clear how *SiMADS51* affects the function of other genes to reduce plant drought tolerance. Even though the expression of *SiMADS51* is significantly induced by exogenous ABA, it is unknown whether its function is dependent on the ABA-mediated pathway. Moreover, functional analysis using knockout mutants is necessary to determine whether *SiMADS51* is required for drought tolerance in foxtail millet. Therefore, further studies will be needed to clarify the mechanism(s) by which *SiMADS51* regulates the response to abiotic stresses.

CONCLUSION

In this study, we report the first systematic analysis of the foxtail millet (*Setaria italica* L.) MADS-box TF family. A total of 72 MADS-box genes including 29 Type I and 43 Type II genes, were identified in the foxtail millet genome. The phylogenetic relationships, chromosomal distribution, conserved motifs, gene structures, and *cis*-acting elements of the 72 foxtail millet MADS-box genes were characterized. Furthermore, the expression patterns of 10 foxtail millet MADS-box genes that are upregulated in response to drought were analyzed in different organs in response to different abiotic stresses. Because the *SiMADS51* gene was found to be strongly induced by drought stress, the function of *SiMADS51* gene was assessed by expression in the model plants *Arabidopsis* and rice (*Oryza sativa* L.). Finally, *SiMADS51* was shown to be involved in the negative regulation of drought stress. Our results provide evidence of the relationship between foxtail millet MADS-box genes and abiotic stresses; therefore, it may be possible to use these genes for both genetically modified crops and in traditional foxtail millet breeding programs.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

D-HM and Y-ZM conceived and designed the experiments. Z-SX edited the manuscript. WZ performed the experiments and wrote the first draft. L-LZ conducted the bioinformatic work and revised the manuscript. LF and H-XP contributed to data analysis. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was financially supported by the National Transgenic Key Project of the Chinese Ministry of Agriculture (2016ZX08002002-010), the National Key Research and Development Program of Wheat Molecular Design and Breeding (2016YFD0101802), and the Programme of Introducing Talents of Innovative Discipline to Universities (Project 111) from the State Administration of Foreign Experts Affairs (#B18042) “Crop breeding for disease resistance and genetic improvement.”

ACKNOWLEDGMENTS

We are grateful to Dr. Xianmin Diao from Crop Science, Chinese Academy of Agricultural Sciences, for providing the foxtail millet (*Setaria italica* L.) seeds.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.659474/full#supplementary-material>

Supplementary Figure 1 | Subcellular localization of *SiMADS*-box proteins in rice protoplasts. Results were visualized with confocal microscopy 16 h after transformation. Scale bars = 10 μ m.

Supplementary Figure 2 | Relative transcript level of *SiMADS51* gene in transgenic *Arabidopsis* and rice lines. **(A)** Relative transcript level of *SiMADS51* gene in 7 transgenic *Arabidopsis* lines. **(B)** Relative transcript level of *SiMADS51* gene in 19 transgenic rice lines.

Supplementary Figure 3 | Overexpression of *SiMADS51* reduces tolerance to drought stress in transgenic rice plants.

Supplementary Table 1 | The 488 MADS-box genes identified from *Arabidopsis*, rice, foxtail millet, potato, and *Setaria viridis* in this study.

Supplementary Table 2 | The amino acid sequences of 20 putative motifs.

Supplementary Table 3 | All the *cis*-regulatory elements of 72 *SiMADS*-box gene promoters predicted by New PLACE online tool.

Supplementary Table 4 | Twenty-five *SiMADS*-box genes that showed a transcriptional response to drought stress were identified from RNA-seq data of drought-treated foxtail millet seedlings.

Supplementary Table 5 | The FPKM (fragments per kilobase of transcript per million mapped reads) values of different tissues of 10 *SiMADS*-box genes.

Supplementary Table 6 | Primers used in this study.

REFERENCES

- Airoidi, C. A., and Davies, B. (2012). Gene duplication and the evolution of plant MADS-box transcription factors. *J. Genet. Genom.* 39, 157–165. doi: 10.1016/j.jgg.2012.02.008
- Alvarez-Buylla, E. R., Liljegren, S. J., Pelaz, S., Gold, S. E., Burgeff, C., Ditta, G. S., et al. (2008). MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. *Plant J.* 24, 457–466. doi: 10.1046/j.1365-3113x.2000.00891.x
- Arora, R., Agarwal, P., Ray, S., Singh, A. K., Singh, V. P., Tyagi, A. K., et al. (2007). MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genomics* 8:242. doi: 10.1186/1471-2164-8-242
- Bailey, T. L., James, J., Grant, C. E., and Noble, W. S. (2015). The MEME suite. *Nucleic Acids Res.* 43, W39–W49. doi: 10.1093/nar/gkv416
- Becker, A., and Theissen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* 29, 464–489. doi: 10.1016/s1055-7903(03)00207-0
- Boden, S. A., and Østergaard, L. (2019). How can developmental biology help feed a growing population? *Development* 146:dev172965. doi: 10.1242/dev.172965
- Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., et al. (2019). MADS-Box genes are key components of genetic regulatory networks involved in abiotic stress and plastic developmental responses in plants. *Front. Plant Sci.* 10:853. doi: 10.3389/fpls.2019.00853
- Changan, S. S., Ali, K., Kumar, V., Garg, N. K., and Tyagi, A. (2018). Absciscic acid biosynthesis under water stress: anomalous behavior of the 9-cis-epoxycarotenoid dioxygenase1 (NCED1) gene in rice. *Biol. Plantarum.* 62, 663–670. doi: 10.1007/s10535-018-0807-2
- Chen, R. G., Ma, J. H., Luo, D., Hou, X. M., Ma, F., Zhang, Y. M., et al. (2019). CaMADS, a MADS-box transcription factor from pepper, plays an important role in the response to cold, salt, and osmotic stress. *Plant Sci.* 280, 164–174. doi: 10.1016/j.plantsci.2018.11.020
- Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-3113x.1998.00343.x
- Day, R. C., Herridge, R. P., Ambrose, B., and Macknight, R. C. (2008). Transcriptome analysis of proliferating *Arabidopsis* endosperm reveals biological implications for the control of syncytial division, cytokinin signaling, and gene expression regulation. *Plant Physiol.* 148, 1964–1984. doi: 10.1105/tpc.110.081737
- De Bodt, S., Raes, J., de Peer, Y. V., and Theissen, G. (2003a). And then there were many: MADS goes genomic. *Trends Plant Sci.* 8, 475–483. doi: 10.1016/j.tplants.2003.09.006
- De Bodt, S., Raes, J., Florquin, K., Rombauts, S., Rouzé, P., Theissen, G., et al. (2003b). Genomewide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. *J. Mol. Evol.* 56, 573–586. doi: 10.1007/s00239-002-2426-x
- Duan, W., Song X., Liu, T., Huang, Z., Ren, J., Hou, X., et al. (2015). Genome-wide analysis of the MADS-box gene family in *Brassica rapa* (Chinese cabbage). *Mol. Genet. Genomics* 290, 239–255. doi: 10.1007/s00438-014-0912-7
- Dubouzet, J. G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E. G., Miura, S., et al. (2003). OsDREB genes in rice, *Oryza sativa* L. encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J.* 33, 751–763. doi: 10.1046/j.1365-3113x.2003.01661.x
- Eddy, S. R., and Pearson, W. R. (2011). Accelerated profile hmm searches. *PLoS Comp. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195
- Fornara, F., Páoenicová, L., Falasca, G., Pelucchi, N., Masiero, S., Ciannamea, S., et al. (2004). Functional characterization of OsMADS18, a member of the API/SQUA subfamily of MADS-box genes. *Plant Physiol.* 135, 2207–2219. doi: 10.1104/pp.104.045039
- Gao, H. H., Wang, Z. M., Li, S. L., Hou, M. L., Zhou, Y., Zhao, Y. Q., et al. (2018). Genome-wide survey of potato MADS-box genes reveals that StMADS1 and StMADS13 are putative downstream targets of tuberigen StSP6A. *BMC Genom.* 19:726. doi: 10.1186/s12864-018-5113-z
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Gopal, S., Park, J. I., Jung, H. J., Ahmed, N. U., Kayum, M. A., Chung, M. Y., et al. (2015). Genome-wide identification and characterization of MADS-box family genes related to organ development and stress resistance in *Brassica rapa*. *BMC Genomics* 16:178. doi: 10.1186/s12864-015-1349-z
- Gramzow, L., Ritz, M. S., and Theissen, G. (2010). On the origin of MADS-domain transcription factors. *Trends Genet.* 26, 149–153. doi: 10.1016/j.tig.2010.01.004
- Gramzow, L., and Theissen, G. (2010). A hitchhiker's guide to the MADS world of plants. *Genome Biol.* 11:214. doi: 10.1186/gb-2010-11-6-214
- Gramzow, L., and Theissen, G. (2015). Phylogenomics reveals surprising sets of essential and dispensable clades of MIKC(c)-group MADS-box genes in flowering plants. *J. Exp. Zool. B. Mol. Dev. Evol.* 324, 353–362. doi: 10.1002/jez.b.22598
- Gu, Z. L., Andre, C., Chen, F. C., Peter, B., and Li, W. H. (2002). Extent of gene duplication in the genomes of drosophila, nematode, and yeast. *Mol. Biol. Evol.* 19, 256–262. doi: 10.1093/oxfordjournals.molbev.a004079
- Gunes, A., Inal, A., Adak, M. S., Bagci, E. G., Cicek, N., and Eraslan, F. (2011). Effect of drought stress implemented at pre- or post-anthesis stage on some physiological parameters as screening criteria in chickpea cultivars. *Russ. J. Plant Physiol.* 55, 59–67. doi: 10.1134/S102144370801007X
- Guo, X. H., Chen, G. P., Cui, B. L., Gao, Q., Guo, J. E., Li, A. Z., et al. (2016). *Solanum lycopersicum* agamous-like MADS-box protein AGL15-like gene, SIMBP11, confers salt stress tolerance. *Mol. Breed.* 36, 125–139.
- He, C. M., Si, C., Silva, J. A. T. D., Li, M. Z., and Duan, J. (2019). Genome-wide identification and classification of MIKC-type MADS-box genes in *Streptophyte* lineages and expression analyses to reveal their role in seed germination of orchid. *BMC Plant Biol.* 19:223. doi: 10.1186/s12870-019-1836-5
- Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Münster, T., and Theissen, G. (2002). Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. *Mol. Biol. Evol.* 19, 801–814. doi: 10.1093/oxfordjournals.molbev.a004137
- Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Res.* 27, 297–300. doi: 10.1093/nar/27.1.297
- Hong, Y. B., Zhang, H. J., Huang, L., Li, D. Y., and Song, F. M. (2016). Overexpression of a stress-Responsive NAC transcription factor gene ONAC022 improves drought and salt tolerance in rice. *Front. Plant Sci.* 7:4. doi: 10.3389/fpls.2016.00004
- Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* 409, 525–529. doi: 10.1038/35054083
- Hu, B., Jin, J., Guo, A. Y., Zhang, H., and Gao, G. (2014). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu187
- Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 18:486. doi: 10.1016/s0168-9525(02)02722-1
- Hwang, S. G., Chen, H. C., Huang, W. Y., Chu, Y. C., Shii, C. T., and Cheng, W. H. (2010). Ectopic expression of rice OsNCED3 in *Arabidopsis* increases ABA level and alters leaf morphology. *Plant Sci.* 178, 12–22. doi: 10.1016/j.plantsci.2009.09.014
- Immink, R. G. H., Gadella, T. W. J., Ferrario, S., Busscher, M., and Angenent, G. C. (2002). Analysis of MADS box protein-protein interactions in living plant cells. *Proc. Natl. Acad. Sci. U S A.* 99, 2416–2421. doi: 10.1073/pnas.042677699
- Irish, V. F., and Amy, A. (2005). Flower development and evolution: gene duplication, diversification and redeployment. *Curr. Opin. Genet. Dev.* 15, 454–460. doi: 10.1016/j.gde.2005.06.001
- Ivica, L., and Peer, B. (2017). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46, D493–D496. doi: 10.1093/nar/gkx922
- Jia, G. H., Huang, X. H., Zhi, H., Zhao, Y., Zhao, Q., Li, W. J., et al. (2013). A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat. Genet.* 45, 957–961. doi: 10.1038/ng.2673
- Jia, J. T., Zhao, P. C., Cheng, L. Q., Yuan, G. X., Yang, W. G., Liu, S., et al. (2018). MADS-box family genes in sheepgrass and their involvement in abiotic stress responses. *BMC Plant Biol.* 18:42. doi: 10.1186/s12870-018-1259-8
- Jin, J. P., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J. C., et al. (2018). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw1328

- Kaufmann, K., Melzer, R., and Theißen, G. (2005). MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347, 183–198. doi: 10.1016/j.gene.2004.12.014
- Khong, G. N., Pati, P. K., Richaud, F., Parizot, B., Bidzinski, P., Mai, C. D., et al. (2015). OsMADS26 negatively regulates resistance to pathogens and drought tolerance in rice. *Plant Physiol.* 169, 2935–2949. doi: 10.1104/pp.15.01192
- Kochhar, S., and Kochhar, V. K. (2005). Expression of antioxidant enzymes and heat shock proteins in relation to combined stress of cadmium and heat in *Vigna mungo* seedlings. *Plant Sci.* 168, 921–929. doi: 10.1016/j.plantsci.2004.11.013
- Krizek, B. A., and Fletcher, J. C. (2005). Molecular mechanisms of flower development: an armchair guide. *Nat. Rev. Genet.* 6, 688–698. doi: 10.1038/nrg1675
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliamet, H. P., et al. (2007). ClustalW and ClustalX version 2. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Lee, S., Choi, S. C., and An, G. (2008b). Rice SVP-group MADS-box proteins, OsMADS22 and OsMADS55, are negative regulators of brassinosteroid responses. *Plant J.* 54, 93–105. doi: 10.1111/j.1365-3113.2008.03406.x
- Lee, S., Woo, Y. M., Ryu, S. I., Shin, Y. D., Kim, W. T., Park, K. Y., et al. (2008a). Further characterization of a rice AGL12 group MADS-box gene, OsMADS26. *Plant Physiol.* 147, 156–168. doi: 10.1104/pp.107.114256
- Li, P., Zhang, Q. L., He, D. N., Zhou, Y., Ni, H. H., Tian, D. G., et al. (2020). AGAMOUS-LIKE67 cooperates with the histone mark reader EBS to modulate seed germination under high temperature. *Plant Physiol.* 184, 529–545. doi: 10.1104/pp.20.00056
- Liu, L. S., White, M. J., and MacRae, T. H. (1999). Transcription factors and their genes in higher plants functional domains, evolution and regulation. *Eur. J. of Biochem.* 262, 247–257. doi: 10.1046/j.1432-1327.1999.00349.x
- Liu, R. H., and Meng, J. L. (2003). MapDraw: a microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data. *Yi Chuan* 25, 317–321. doi: 10.3321/j.issn:0253-9772.2003.03.019
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Lu, S. N., Wang, J. Y., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2019). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 48, D265–D268. doi: 10.1093/nar/gkz991
- Masiero, S., Colombo, L., Grini, P. E., Schmittgen, A., and Kater, M. M. (2011). The emerging importance of type I MADS box transcription factors for plant reproduction. *Plant Cell* 23, 865–872. doi: 10.1105/tpc.110.081737
- McGonigle, B., Bouhidel, K., and Irish, V. (1996). Nuclear localization of the Arabidopsis APETALA3 and PISTILLATA homeotic gene products depends on their simultaneous expression. *Genes Dev.* 10, 1812–1821. doi: 10.1101/gad.10.14.1812
- Moser, M., Asquini, E., Miolli, G. V., Weigl, K., and Si-Ammour, A. (2020). The MADS-Box gene MdDAM1 controls growth cessation and bud dormancy in Apple. *Front. Plant Sci.* 11:1003. doi: 10.3389/fpls.2020.01003
- Nardeli, S. M., Artico, S., Aoyagi, G. M., de Moura, S. M., da Franca, S. T., Grossi-de-Sa, M. F., et al. (2018). Genome-wide analysis of the MADS-box gene family in polyploid cotton (*Gossypium hirsutum*) and in its diploid parental species (*Gossypium arboreum* and *Gossypium raimondii*). *Plant Physiol. Biochem.* 127, 169–184. doi: 10.1016/j.plaphy.2018.03.019
- Parenicová, L., Folter, S. D., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., et al. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *Plant Cell* 15, 1538–1851. doi: 10.1105/tpc.011544
- Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., et al. (2014). Structural basis for the oligomerization of the MADS domain transcription factor SEPALLATA3 in *Arabidopsis*. *Plant Cell* 26, 3603–3615. doi: 10.1105/tpc.114.127910
- Riechmann, J. L., and Ratcliffe, O. J. (2000). A genomic perspective on plant transcription factors. *Curr. Opin. Plant Biol.* 3, 423–434. doi: 10.1016/S1369-5266(00)00107-2
- Riechmann, J. L., Wang, M., and Meyerowitz, E. M. (1996). DNA-binding properties of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. *Nucleic Acids Res.* 24, 3134–3141. doi: 10.1016/S0025-7125(05)70504-6
- Sallaud, C., Meynard, D., Bostel, J. V., Bès, C. G. M., Brizard, J. P., Larmande, P., et al. (2003). Guiderdoni Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theoretical Appl. Genet.* 106, 1396–1408. doi: 10.1007/s00122-002-1184-x
- Schilling, S., Kennedy, A., Pan, S., Jermin, L. S., and Melzer, R. (2020). Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. *New Phytol.* 22, 511–529. doi: 10.1111/nph.16122
- Schilling, S., Pan, S., Kennedy, A., and Melzer, R. (2018). MADS-box genes and crop domestication: the jack of all traits. *J. Exp. Bot.* 69, 1447–1469. doi: 10.1093/jxb/erx479
- Shah, K., and Nahakpam, S. (2012). Heat exposure alters the expression of SOD, POD, APX and CAT isozymes and mitigates low cadmium toxicity in seedlings of sensitive and tolerant rice cultivars. *Plant Physiol. Biochem.* 57, 106–113. doi: 10.1016/j.plaphy.2012.05.007
- Singh, A., Giri, J., Kapoor, S., Tyagi, A. K., and Pandey, G. K. (2010). Protein phosphatase complement in rice: genome-wide identification and transcriptional analysis under abiotic stress conditions and reproductive development. *BMC Genomics* 11:435. doi: 10.1186/1471-2164-11-435
- Singh, K. B., Foley, R. C., and Oate-Sánchez, L. (2002). Transcription factors in plant defense and stress responses. *Curr. Opin. Plant Biol.* 5, 430–436. doi: 10.1016/S1369-5266(02)00289-3
- Sudhir, K., Glen, S., and Koichiro, T. (2016). MEGA7: molecular evolutionary genetics analysis Version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Tapia-López, R., García-Ponce, B., Dubrovsky, J. G., Garay-Arroyo, A., Pérez-Ruiz, R. V., Kim, S. H., et al. (2008). An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in *Arabidopsis*. *Plant Physiol.* 146, 1182–1192. doi: 10.1104/pp.107.108647
- Theißen, G., Melzer, R., and Rümpler, F. (2016). MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development* 143, 3259–3271. doi: 10.1242/dev.134080
- Theißen, G., Rümpler, F., and Gramzow, L. (2018). Array of MADS-box genes: facilitator for rapid adaptation? *Trends Plant Sci.* 23, 563–576. doi: 10.1016/j.tplants.2018.04.008
- Tiwari, S., Spielman, M., Schulz, R., Oakey, R. J., Kelsey, G., Salazar, A., et al. (2010). Transcriptional profiles underlying parent-of-origin effects in seeds of *Arabidopsis thaliana*. *BMC Plant Biol.* 10:72. doi: 10.1186/1471-2229-10-72
- Walia, H., Josefsson, C., Dilkes, B., Kirkbride, R., Harada, J., and Comai, L. (2009). Dosage-dependent deregulation of an AGAMOUSLIKE gene cluster contributes to interspecific incompatibility. *Curr. Biol.* 19, 1128–1132. doi: 10.1016/j.cub.2009.05.068
- Wang, Y. S., Zhang, J. L., Hu, Z. L., Guo, X. H., Tian, S. B., and Chen, G. P. (2019). Genome-wide analysis of the MADS-box transcription factor family in *Solanum lycopersicum*. *Int. J. Mol. Sci.* 20:2961. doi: 10.3390/ijms20122961
- Wang, Z., Wang, F. X., Hong, Y. C., Yao, J. J., Ren, Z. Z., Shi, H. Z., et al. (2018). The flowering repressor SVP confers drought resistance in *Arabidopsis* by regulating abscisic acid catabolism. *Mol. Plant* 011, 1184–1197. doi: 10.1016/j.molp.2018.06.009
- Wei, M. M., Wang, Y. J., Pan, R. R., and Li, W. G. (2018). Genome-wide identification and characterization of MADS-box family genes related to floral organ development and stress resistance in *Hevea brasiliensis* Müll. *Arg. Forests* 9:304. doi: 10.3390/f9060304
- Wei, X., Wang, L., Yu, J., Zhang, Y., Li, D., and Zhang, X. (2015). Genome-wide identification analysis of the MADS-box gene family in sesame. *Gene* 569, 66–76. doi: 10.1016/j.gene.2015.05.018
- Wu, R. M., Wang, T. C., Warren, B. A. W., Thomson, S. J., Allan, A. C., Macknight, R. C., et al. (2018). Kiwifruit SVP2 controls developmental and drought-stress pathways. *Plant Mol. Biol.* 96, 233–244. doi: 10.1007/s11103-017-0688-3
- Wuest, S. E., Vijverberg, K., Schmidt, A., Weiss, M., Gheyselsinck, J., Lohr, M., et al. (2010). Arabidopsis female gametophyte gene expression map reveals similarities between plant and animal gametes. *Curr. Biol.* 20, 506–512. doi: 10.1016/j.cub.2010.01.051
- Xiao, B., Huang, Y. M., Tang, N., and Xiong, L. Z. (2007). Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theor. Appl. Genet.* 115, 35–46. doi: 10.1007/s00122-007-0538-9

- Xiong, H. Y., Li, J. J., Liu, P. L., Duan, J. Z., Zhao, Y., Guo, X., et al. (2014). Overexpression of OsMYB48-1, a Novel MYB-Related transcription factor, enhances drought and salinity tolerance in rice. *PLoS One* 9:e92913. doi: 10.1371/journal.pone.0092913
- Yang, A., Dai, X. Y., and Zhang, W. H. (2012). A R2R3-type MYB gene, OsMYB2, is involved in salt, cold, and dehydration tolerance in rice. *J. Exp. Bot.* 63, 2541–2556. doi: 10.1093/jxb/err431
- Yang, Y., and Jack, T. (2004). Defining subdomains of the K domain important for protein-protein interactions of plant MADS proteins. *Plant Mol. Biol.* 55, 45–59. doi: 10.1007/s11103-004-0416-7
- Yu, T. F., Zhao, W. Y., Fu, J. D., Liu, Y. W., Chen, M., Zhou, Y. B., et al. (2018). Genome-wide analysis of CDPK family in foxtail millet and determination of SiCDPK24 functions in drought Stress. *Front. Plant Sci.* 9:651. doi: 10.3389/fpls.2018.00651
- Zhang, G. Y., Liu, X., Quan, Z. W., Cheng, S. F., Xu, X., Pan, S. K., et al. (2012). Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat. Biotechnol.* 30, 549–554. doi: 10.1038/nbt.2195
- Zhang, H. K., Gao, S. H., Martin, J. L., Hu, S. N., and Chen, W. H. (2012). EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucl. Acids Res.* 40, W569–W572. doi: 10.1093/nar/gks576
- Zhang, Y., Su, J. B., Duan, S., Ao, Y., and Wang, H. B. (2011). A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods* 7:30. doi: 10.1186/1746-4811-7-30
- Zhang, Y. T., Tang, D. Q., Lin, X. C., Ding, M. Q., and Tong, Z. K. (2018). Genome-wide identification of mads-box family genes in moso bamboo (*Phyllostachys edulis*) and a functional analysis of PeMADS5 in flowering. *BMC Plant Biol.* 18:176. doi: 10.1186/s12870-018-1394-2
- Zhang, Z., Li, H., Zhang, D., Liu, Y., Fu, J., Shi, Y., et al. (2012). Characterization and expression analysis of six MADS-box genes in maize (*Zea mays* L.). *J. Plant Physiol.* 169, 797–806. doi: 10.1016/j.jplph.2011.12.020
- Zhang, Z., Li, J., Zhao, X. Q., Wang, J., Wong, G. K., and Yu, J. (2006). KaKs calculator: calculating Ka and Ks through model selection and model averaging. *Genom. Proteom. Bioinform.* 4, 259–263. doi: 10.1016/S1672-0229(07)60007-2
- Zhao, S. P., Xu, Z. S., Zheng, W. J., Zhao, W., Wang, Y. X., Yu, T. F., et al. (2017). Genome-wide analysis of the RAV family in soybean and functional identification of GmRAV-03 involvement in salt and drought stresses and exogenous ABA treatment. *Front. Plant Sci.* 8:905. doi: 10.3389/fpls.2017.00905
- Zhao, T., Ni, Z., Dai, Y., Yao, Y., Nie, X., and Sun, Q. (2006). Characterization and expression of 42 MADS-box genes in wheat (*Triticum aestivum* L.). *Mol. Genet. Genom.* 276, 334–350. doi: 10.1007/s00438-006-0147-3
- Zhao, W., Liu, Y. W., Zhou, J. M., Zhao, S. P., Zhang, X. H., and Min, D. H. (2016). Genome-wide analysis of the lectin receptor-like kinase family in foxtail millet (*Setaria italica* L.). *Plant Cell Tiss. Organ. Cult.* 127, 335–346. doi: 10.1007/s11240-016-1053-y
- Zhao, Y., Li, X. Y., Chen, W. J., Peng, X. J., Cheng, X., Zhu, S. W., et al. (2011). Whole-genome survey and characterization of MADS-box gene family in maize and sorghum. *Plant Cell Tiss. Organ. Cult.* 105, 159–173. doi: 10.1007/s11240-010-9848-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhao, Zhang, Xu, Fu, Pang, Ma and Min. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Variation for Photoperiod and Temperature Sensitivity in the Global Mini Core Collection of Sorghum

Hari D. Upadhyaya^{1,2*}, M. Vetriventhan¹ and Vania C. R. Azevedo¹

¹ Genebank, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, ² Plant Genome Mapping Laboratory, University of Georgia, Athens, GA, United States

OPEN ACCESS

Edited by:

Wolfram Weckwerth,
University of Vienna, Austria

Reviewed by:

Zhenbin Hu,
Kansas State University, United States
Marcello Mastrorilli,
Council for Agricultural Research
and Economics (CREA), Italy

*Correspondence:

Hari D. Upadhyaya
harideo.upadhyaya@gmail.com;
hari.upadhyaya@uga.edu

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 12 June 2020

Accepted: 17 May 2021

Published: 29 June 2021

Citation:

Upadhyaya HD, Vetriventhan M
and Azevedo VCR (2021) Variation
for Photoperiod and Temperature
Sensitivity in the Global Mini Core
Collection of Sorghum.
Front. Plant Sci. 12:571243.
doi: 10.3389/fpls.2021.571243

Information on photoperiod and temperature sensitivity of sorghum germplasm is important to identify appropriate sources for developing cultivars with a broad adaptation. The sorghum mini core collection consisting of 242 accessions along with three control cultivars were evaluated for days to 50% flowering (DFL) and plant height in two long-day rainy and two short-day post-rainy seasons, and for grain yield and 100-seed weight in the two post-rainy seasons. Differences in DFL and cumulative growing degree days (CGDD) in the rainy and post-rainy seasons were used to classify the accessions for photoperiod and temperature sensitivity. Results revealed 18 mini core landraces as photoperiod and temperature insensitive (PTINS), 205 as photoperiod sensitive and temperature insensitive (PSTINS), and 19 as photoperiod and temperature-sensitive (PTS) sources. The 19 PTS sources and 80 PSTINS sources took less DFL in the long-day rainy seasons than in the short-day post-rainy season indicating their adaptation to the rainy season and a possible different mechanism than that trigger flowering in the short-day sorghums. In all three groups, several accessions with desirable combinations of agronomic traits were identified for use in the breeding programs to develop climate-resilient cultivars and for genomic studies to identify genes responsible for the photoperiod and temperature responses.

Keywords: sorghum, photosensitivity, germplasm, photoperiod insensitive, cumulative growing degree days

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is an important multipurpose crop grown for food, feed, fodder and fuel purposes. It is a C₄ crop with high photosynthetic efficiency, cultivated in both temperate and tropical regions between 40°N to 40°S of the equator (Doggett, 1988). Sorghum ranks fifth among the cereal crops after maize (*Zea mays* L.), rice (*Oryza sativa* L.), wheat (*Triticum aestivum* L.), and barley (*Hordeum vulgare* L.). During 2018, sorghum was grown on an estimated area of 42.14 Mha with a production of 59.34 Mt in over 100 countries, with a productivity of 1408 kg ha⁻¹. Globally, sorghum is largely cultivated during the rainy season in diverse conditions which is often affected by moisture stress and also several diseases and insect pests affect crop productivity (Assefa et al., 2010; Upadhyaya et al., 2019). In India sorghum is grown in two distinct seasons: kharif (rainy season, June–October) and rabi (post-rainy season, October–February). Kharif season is characterized by long-days with high temperature and rabi season by the short-days with low temperature. It is therefore important to develop high-yielding cultivars resistant to various biotic and abiotic stresses and adapted to diverse climatic conditions.

Variation in the germplasm collection is key to crop improvement. Vast germplasm collections are conserved in the national and international Genebanks. Globally over 23,600 accessions of sorghum have been conserved in the genebanks having a rich diversity for various traits of economic importance (Upadhyaya et al., 2016b). The ICRISAT Genebank has the largest collection of over 41,000 sorghum accessions from 93 countries. Large size of germplasm collection and non-availability of reliable data on traits of economic importance restricted breeders from using germplasm collections and caused them to use their working collections (Upadhyaya et al., 2014d). This results in low use of germplasm in crop improvement and consequently a narrow genetic base of crop cultivars (Dalrymple, 1986; Vellve, 1992; Upadhyaya et al., 2014b). Genetic studies indicated that the sorghum wild and weedy relatives and landraces are more diverse than the improved hybrids (Muraya et al., 2011; Mace et al., 2013). For example, in the United States, the improved sweet sorghum lines were mostly derived from six landraces therefore the majority of the lines have similar genes/alleles for high Brix % (Murray et al., 2009). The genetic diversity assessment sorghum hybrids widely cultivated in the United States between 1980 and 2008 indicated that the hybrids released during the 2000s had the least number of new alleles as compared to the previous two decades (Smith et al., 2010). Similarly, a decline in genetic diversity was reported in Australian sorghum breeding programs as a result of a strong selection for resistance to the sorghum midge, stay-green (a drought resistance trait) and other agronomic traits (Jordan et al., 1998, 2003). In sorghum which is a short-day plant (SDP), several germplasm accessions do not flower or take much more time in the long-day rainy season and thus restrict the use of germplasm in crop breeding programs.

The germplasm diversity representative subsets such as core collection (10% of the entire collection) and mini core collection (10% of core collection or 1% of the entire collection) conserved in a genebank are the ideal genetic resources for identifying new sources of variation for traits of interest (Upadhyaya and Ortiz, 2001; Upadhyaya et al., 2013, 2014c). In this study, we evaluated 242 accessions of sorghum mini core collection (Upadhyaya et al., 2009) for photoperiod and temperature sensitivity in the two rainy and two post-rainy seasons at ICRISAT, Patancheru, India, to identify insensitive accessions that can be used in crop improvement to develop high-yielding photoperiod and temperature insensitive cultivars with a broad genetic base and greater adaptation.

MATERIALS AND METHODS

Description of Materials

The 242 accessions in the sorghum mini core collection (Upadhyaya et al., 2009) and three controls (IS 33844, IS 18758 and IS 2205) were utilized in this study. The mini core accessions originating (collection site) from 57 countries, and it included all the five basic races (caudatum 16.1%, durra 12.4%, guinea 12%, kafir 8.7%, and bicolor 8.3%) and 10 intermediate races (caudatum-bicolor 12.4%; guinea-caudatum 11.2%; durra-caudatum 7.9%; durra-bicolor and kafir-caudatum each 2.9%;

kafir-durra 1.7%; guinea-kafir 1.2%; and guinea-bicolor, guinea-durra, and kafir-bicolor each 0.8%) of sorghum (Upadhyaya et al., 2009). The two controls, IS 33844 and IS 18758 are the germplasm accessions and selection from these were released as cultivars. IS 33844 was released as Parbhani Moti in India, which is the most popular drought-tolerant cultivar grown under receding soil moisture conditions during the post-rainy seasons and possesses excellent grain quality attributes (Upadhyaya et al., 2017). The selection from IS 18758 was released in Burkina Faso as E-35-1 and in Burundi as Gambella 1107 (Upadhyaya et al., 2014a). IS 2205 is a durra-bicolor landrace originating from India is resistant to shoot fly [*Atherigona soccata* (Rondani)] and stem borer [*Chilo partellus* (Swinhoe)] (Sharma et al., 2003).

Experimental Details

The experiments were conducted at the ICRISAT, Patancheru (17.51° N, 78.27° E, and 545 masl), India during the 2010 and 2011 rainy and 2010–2011 and 2011–2012 post-rainy seasons. The precision fields of the ICRISAT, Patancheru center having uniform fertility with a gentle slope of 0.5% were used. The experiments were conducted in an Alpha design in the rainy seasons with three replications. In the post-rainy seasons, accessions were planted in a split-plot design in three replications using drought stress and optimally irrigated treatments as the main plot and genotypes as the subplots in five maturity groups (Upadhyaya et al., 2017), and data from the optimally irrigated condition was used for this study. The experimental materials were planted in the last week of June in the rainy season and in the second week of October in the post-rainy season each year. Each accession was sown in a single row of 4 m long plot and the spacing between plots was 75 cm. Seeds were sown at a uniform depth of 2–3 cm using a tractor-mounted four-cone planter, and excess seedlings were thinned during 2 to 3rd weeks after sowing, by maintaining the plant-to-plant spacing of about 10 cm. The crop-specific agronomic practices such as weeding, thinning, irrigation and plant protection measures were followed. A basal dose of Ammonium phosphate was applied at the rate of 150 kg ha⁻¹, and urea was applied as topdressing at the rate of 100 kg ha⁻¹ during the 3 weeks after planting. A ridge and furrow system of cultivation was adopted, and each time, the experimental plots received about 7 cm of irrigation water. Observations on days to 50% flowering (DFL, the day when 50% or more of the plants had reached anthesis in a plot) were recorded in each of the two rainy and two post-rainy seasons. Plant height was recorded on five representative plants at maturity in the rainy and post-rainy seasons whereas data on grain yield and 100-seed weight was recorded only in the post-rainy season on five representative plants.

Data Analysis

Data on daily minimum and maximum temperatures was obtained from the ICRISAT intranet¹ and was used to calculate the growing degree days (GDD) in the two rainy and two post-rainy seasons separately.

¹<http://intranet.icrisat.org/IcrisatLocal/weather.aspx>

The GDDs were computed using the following equation:

$$\text{GDD} = [(\text{daily maximum temperature} + \text{daily minimum temperature})/2] - 10^{\circ}\text{C, the base temperature for sorghum}].$$

The cumulative growing degree days (CGDD) were computed as follows:

CGDD = the sum of all daily GDD from the effective date of sowing to the date of 50% flowering.

For sorghum a base temperature of 10°C (Gerik et al., 2003) was used. Data on DFL, CGDD, and plant height in the two rainy and two post-rainy and grain yield and 100-seed weight in the two post-rainy seasons were analyzed separately using Genstat 17² following residual (or restricted) maximum likelihood (REML) (Patterson and Thompson, 1971). Combined analyses of two rainy seasons and two post-rainy seasons were performed separately considering genotypes as random and seasons fixed. Variance components due to genotypes and genotype \times environment (season) were estimated and tested for their significance against their respective standard errors. Wald's statistic (Wald, 1943) for the season was calculated and tested to determine the significance of the two rainy or two post-rainy seasons. For DFL, CGDD, plant height, grain yield and 100-seed weight, BLUPs (balanced linear unbiased predictors) were obtained for combined analysis of two rainy and two post-rainy seasons, separately.

Classification of Mini Core Accessions for Photoperiod and Temperature Sensitivity

The mini core accessions were categorized into photoperiod and temperature-sensitive/insensitive considering the differences in DFL and CGDD requirement during the long-day rainy season and short-day post-rainy seasons. When the difference between DFL in rainy and post-rainy seasons was equal to twice the standard error of the mean (± 4 calendar days), the accessions were classified as photoperiod-insensitive while the remaining were considered as photoperiod-sensitive. Similarly, in the case of temperature sensitivity, the accessions were classified as temperature-sensitive when the difference in CGDD requirements in the rainy and post-rainy seasons was equivalent to ± 4 calendar days (-57 to $+57^{\circ}\text{d}$) while the remaining accessions on both sides (+ and -) were categorized as temperature-insensitive. Finally, both photoperiod and temperatures sensitivities were considered together which resulted in three groups of accessions: photoperiod- and temperature-insensitive (PTINS), photoperiod-sensitive and temperature-insensitive (PSTINS), and photoperiod- and temperature-sensitive (PTS).

The phenotypic correlations among DFL, CGDD, and plant height in both the rainy and post-rainy seasons and grain yield and 100-seed weight in the post-rainy season were calculated using Genstat 17 (see Text Footnote 2) to determine the

relationships between traits in the entire mini core collection, PTINS, PSTINS, and PTS groups of accessions.

RESULTS

Variance Components, Mean and Range

The REML analysis of individual rainy and post-rainy season data, and combined of two rainy and two post-rainy seasons revealed significant genotypic variance (σ^2_g) for DFL, CGDD, plant height, grain yield and 100-seed weight (**Supplementary Table 1**). The significant interactions between genotypes and seasons (σ^2_{ge}) was observed for all the traits in both rainy and post-rainy seasons, except in group 5 where the plant height, grain yield per plant and 100-seed weight were not significant (**Supplementary Table 1**). Wald's statistics was significant for all the traits in the rainy and post-rainy seasons except in the G5 in the post-rainy season for 100-seed weight.

In the mini core collection, the DFL ranged from 45 to 152 days in the 2010 rainy season and 42–142 days in the 2011 rainy season (**Table 1**). On average, mini core accessions flowered 10 days later in 2010 than in the 2011 rainy season indicating differences in the two rainy seasons mainly due to varying GDD per calendar day during flowering, besides other factors including the genotype by environment interactions. Averaged over two rainy seasons the DFL ranged from 44 (IS 12706) to 148 days (IS 15744) (**Table 1** and **Supplementary Table 2**). The differences in the DFL have also reflected the differences in the CGDD in the two rainy seasons. The CGDD to flower ranged from 758 to $2,403^{\circ}\text{d}$ in the 2010 rainy season with a greater mean ($1,415^{\circ}\text{d}$) than in the 2011 rainy season (range 731 – $2,275^{\circ}\text{d}$, mean $1,297^{\circ}\text{d}$). Averaged over both the rainy seasons the CGDD varied from 756 (IS 12706) to $2,356$ (IS 15744) with a mean value of $1,369^{\circ}\text{d}$ (**Supplementary Table 2**). In the post-rainy seasons also the DFL differed between 2010–2011 and 2011–2012 seasons. The DFL ranged from 47 to 109 in the 2010–2011 and 51 to 132 in the 2011–2012 post-rainy season. On average mini core accessions took 5-days more to flower in the 2011–2012 season (75 days) than in the 2010–2011 season (70 days). Averaged over both the post-rainy seasons, DFL ranged from 51 days (IS 32245) to 120 (IS 31446) with an average of 73 days (**Supplementary Table 2**). The CGDD required to flower in the 2010–2011 post-rainy season were less (793°d) with a lower range (577 to $1,270^{\circ}\text{d}$) than in the 2011–2012 post-rainy season (mean 966°d , range 693 to $1,698^{\circ}\text{d}$). This indicated that the two post-rainy seasons affected the expression of flowering in the mini core collection accessions. On average over the two post-rainy seasons, the mini core accessions took 880°d to flower with IS 32245 flowering earliest requiring only 658°d and IS 31446 being the latest requiring $1,485^{\circ}\text{d}$. Both in terms of DFL and the CGDD requirements highly significant genotype \times season (year) interaction was observed (**Supplementary Table 1**). Plant height of mini core accessions also showed a great deal of differential expression in the two rainy and two post-rainy seasons. On average plant height was higher in the 2011 rainy season (326 cm) than in the 2010 rainy season (281 cm) with a greater range in the former (117–501 cm) than the latter

²<https://www.vsnl.co.uk/>

(126–406 cm) (**Table 1**). Similarly, in the 2011–2012 post-rainy season mean (256 cm) and ranges (99–390 cm) for plant heights were more than in the 2010–2011 post-rainy season (214 and 105–347 cm) (**Table 1**).

Photoperiod and Temperature-Sensitive and Insensitive Mini Core Accessions

The criteria of twice the standard error of the mean for differences in DFL and CGDD requirements between the long-day rainy seasons (combined of two rainy seasons) and the short-day post-rainy seasons (combined of two post-rainy), the mini core collection accessions were classified into three groups: PTINS (18 accessions), PSTINS (205 accessions) and PTS (19 accessions) (**Supplementary Table 2**). The three control cultivars were also grouped into PTINS (IS 2205 and IS 33844) and PSTINS (IS 18758). On average, the PTINS took 57–84 days in the rainy seasons and 56–87 days in the post-rainy seasons and the differences in DFL of the 18 individual PTINS accessions between the rainy and post-rainy seasons were within ± 4 days (**Table 2**) indicating their adaptation to both rainy and post-rainy seasons. The CGDD requirements of the PTINS accessions were more in the rainy season (1,159°d) than in the post-rainy season (837°d). For 18 PTINS mini core accessions individually, the CGDD requirement ranged from 971 to 1,386°d in the rainy seasons and 711 to 1,041°d in the post-rainy season and differences between two seasons ranged from 218°d (IS 4060) to 417°d (IS 7131). The two PTINS controls, IS 2205 and IS 33844 required to flower 1,267 and 1,239°d in the rainy season and 960 and 924°d in the post-rainy season, respectively.

The PTS landraces flowered early in the rainy seasons and on average took 52 days compared to 72 days in the post-rainy season and showed similar CGDD requirement ($\pm 57^\circ\text{d}$). The DFL requirement of the 19 individual PTS accessions ranged from 46 to 66 days in the rainy season and 60 to 94 days in the post-rainy season (**Table 3**). The CGDD requirement was between 785 and 1,110°d in the rainy season and 738 and 1,121°d in the post-rainy season.

The 205 PSTINS mini core accessions on average took 14 days and 548°d more in the rainy season (87 and 1,433°d) for flowering than in the post-rainy season (73 and 885°d) (**Table 4**). However, within these 205 mini core accessions, there were vast differences in terms of DFL and CGDD requirements in the rainy season *vis-a-vis* their requirements in the post-rainy season. The DFL ranged from 44 to 148 days in the rainy and from 51 to 120 days in the post-rainy season and the CGDD requirement from 756 to 2,356 °d in the rainy and from 658 to 1,485°d in the post-rainy season. This indicated different adaptation patterns of PSTINS accessions in terms of DFL and CGDD requirement for flowering. This group was therefore further categorized into three sub-groups. The accessions that took fewer days and fewer CGDD for flowering in the rainy season than in the post-rainy season were grouped into subgroup 1 following (Upadhyaya et al., 2018). None of the mini core accessions fell into this subgroup. The accessions that took fewer days and more CGDD in the rainy season than in the post-rainy accessions were grouped into Subgroup 2 and the accessions, which took more days and more CGDD in the rainy season than in the post-rainy were classified into Sub-group 3. In the mini core collection, 80 accessions were categorized into Subgroup 2 and 125 and control IS 18758 into Subgroup 3. The accessions in the Subgroup 2 on average took 11 days less in the rainy season (60 days) than in the post-rainy season (71 days) whereas those in the Subgroup 3 took 30 days more in the rainy season (104 days) than in the post-rainy season (74 days). Individually the difference between DFL in the rainy season and the post-rainy season ranged from –22 to –5 days in subgroup 2 compared to 5–74 days in the Subgroup 3. Of the 125 accessions in the Subgroup 3, 59 accessions took on average 30–74 days more, 24 took 45–74 days more and 7 took 61–74 days more in the rainy season than in the post-rainy season (**Supplementary Table 2**). Contrastingly, in Subgroup 2, 20 accessions took 15–22 days less in the rainy season than in the post-rainy season, indicating the differential response of these two subgroups for flowering. These accessions could be useful sources for the sorghum improvement program for the two distinct conditions,

TABLE 1 | Mean and range of different traits in the 2010 and 2011 rainy and 2010–2011 and 2011–2012 post-rainy seasons and combined of two rainy and two post-rainy seasons at ICRISAT, Patancheru, India.

Trait	Rainy 2010		Rainy 2011		Combined	
	Mean	Range	Mean	Range	Mean	Range
Days to 50% flowering (days)	86.6	45–152	77.1	42–142	83	44–148
Cumulative growing degree days (°d)	1415	758–2403	1297	731–2275	1369	756–2356
Plant height (cm)	281	126–406	326	117–501	301	115–446

Trait	Post-rainy 2010–11		Post-rainy 2011–12		Combined	
	Mean	Range	Mean	Range	Mean	Range
Days to 50% flowering (days)	70	47–109	75	51–132	73	51–120
Cumulative growing degree days (°d)	793	577–1270	966	693–1698	880	658–1485
Plant height (cm)	214	105–347	256	99–390	235	106–369
Grain yield per plant (g)	21	2–41	26	3–74	24	2–54
100-seed weight (g)	2.50	0.50–5.00	2.40	0.50–5.70	2.43	0.48–5.33

TABLE 2 | Performance of photoperiod and temperature insensitive landraces identified in the sorghum mini core collection, evaluated during two rainy and two post-rainy seasons at ICRISAT, Patancheru, India.

IS No.	Origin	Race	Days to 50% flowering		Cumulative growing degree days (CGDD)		Plant height (cm)		Grain yield (g)	100-seed weight (g)
			Rainy	Post-rainy	Rainy	Post-rainy	Rainy	Post-rainy		
473	United States	GK	64	66	1074	799	269	211	14	1.48
2864	South Africa	C	61	58	1025	728	189	137	20	2.14
4060	India	DB	57	61	971	753	274	194	16	2.09
4581	India	D	72	75	1201	893	286	226	33	2.63
4631	India	D	74	72	1233	857	309	244	31	2.89
5094	India	D	74	76	1223	905	335	272	30	2.45
7131	Uganda	DC	84	81	1386	969	299	231	27	2.24
11619	Ethiopia	DB	83	87	1369	1041	301	331	20	1.87
12697	Australia	B	76	72	1268	869	335	260	10	1.43
12937	Ethiopia	K	69	69	1151	830	331	245	24	2.04
14290	Botswana	KD	67	69	1120	833	297	238	25	2.41
17941	India	C	60	57	1009	717	185	129	16	2.39
20743	United States	B	67	63	1122	770	326	231	17	1.92
21645	Malawi	GK	68	65	1141	791	302	219	11	1.70
24348	India	C	59	56	1006	711	201	153	17	2.19
29239	Swaziland	K	68	71	1140	856	214	193	26	2.99
29568	Lesotho	KC	63	67	1069	806	292	237	29	2.55
29714	Zimbabwe	KD	82	78	1350	941	263	221	26	1.77
Mean			69	69	1159	837	278	221	22	2.20
Range			57–84	56–87	971–1386	711–1041	185–335	129–331	10–33	1.43–2.99
Controls										
2205	India	DB	76	80	1267	960	334	251	26	2.10
18758	Ethiopia	GC	78	69	1288	833	283	166	29	2.40
33844	India	D	75	77	1239	924	325	262	41	3.60
SE(d)			2.45	1.48	39.88	14.18	16.00	11.16	4.83	0.19
LSD ($P \leq 0.05$)			4.80	2.89	78.17	27.79	31.36	21.87	9.46	0.37
CV (%)			4.18	1.95	4.11	1.53	7.51	4.63	21.45	8.10

TABLE 3 | Performance of photoperiod and temperature sensitive landraces identified in the sorghum mini core collection evaluated during two rainy and two post-rainy seasons at ICRISAT, Patancheru, India.

IS No.	Origin	Race	Days to 50% flowering		Cumulative growing degree days		Plant height (cm)		Grain yield (g)	100-seed weight (g)
			Rainy	Post-rainy	Rainy	Post-rainy	Rainy	Post-rainy		
1212	China	KB	47	65	801	788	255	238	17	2.38
1219	China	GB	47	64	803	779	261	241	22	2.00
16151	Cameroon	CB	48	68	819	816	238	242	22	1.64
19676	Zimbabwe	K	66	88	1106	1048	189	156	20	2.50
20727	United States	B	46	65	785	791	221	219	14	1.37
22616	Myanmar	B	62	88	1052	1054	293	287	15	1.33
24503	South Africa	B	49	73	839	880	259	248	2	0.48
26701	South Africa	CB	61	87	1031	1039	245	224	20	1.36
28614	Yemen	DC	49	66	828	798	282	240	18	2.15
29314	Swaziland	DC	66	92	1109	1097	298	278	22	1.87
29689	Zimbabwe	K	66	94	1110	1121	281	271	25	2.27
30383	China	CB	48	64	819	776	266	247	17	2.06
30400	China	CB	49	66	830	803	286	271	22	2.05
30417	China	CB	49	64	833	782	288	259	25	2.33
30450	China	CB	47	64	803	776	231	202	18	2.43
30451	China	CB	49	65	839	786	255	223	28	2.11
30466	China	CB	46	60	785	738	248	240	18	2.01
30507	Korea	CB	49	64	828	776	272	235	21	2.04
30508	Korea	CB	49	66	828	801	214	219	19	1.92
Control										
2205	India	DB	76	80	1267	960	334	251	26	2.14
18758	Ethiopia	GC	78	69	1288	833	283	166	29	2.37
33844	India	D	75	77	1239	924	325	262	41	3.60
Mean			52	72	887	866	257	239	19	1.91
Range			46–66	60–94	785–1110	738–1121	189–298	156–287	2–28	0.48–2.50
SE(d)			2.45	1.48	39.88	14.18	16	11.16	4.83	0.19
LSD ($P \leq 0.05$)			4.8	2.89	78.17	27.79	31.36	21.87	9.46	0.37
CV (%)			4.18	1.95	4.11	1.53	7.51	4.63	21.45	8.10

TABLE 4 | Performance of 15 best yielding landraces identified in the photoperiod sensitive and temperature insensitive subgroups 2 and 3 in the sorghum mini core collection evaluated in two rainy and two post-rainy seasons at ICRISAT, Patancheru, India.

IS No.	Origin	Race	Days to 50% flowering		Cumulative growing degree days		Plant height (cm)		Grain yield (g)	100-seed weight (g)
			Rainy	Post-rainy	Rainy	Post-rainy	Rainy	Post-rainy		
Sub-group 2 (80 accessions)										
4698	India	D	70	78	1175	935	282	230	54	3.36
22294	Botswana	K	67	77	1128	926	298	285	36	3.12
29241	Swaziland	KC	59	73	1007	874	266	273	36	2.35
29187	Swaziland	GC	65	79	1093	951	278	256	35	2.11
13971	South Africa	C	65	77	1089	924	282	264	35	2.23
4515	India	D	70	75	1162	891	311	244	33	3.15
29627	South Africa	DC	56	69	952	827	288	249	33	2.27
19450	Botswana	GK	57	66	971	798	237	198	32	2.47
29606	South Africa	K	62	77	1053	937	310	264	31	2.96
29358	Lesotho	K	62	68	1044	816	249	202	30	2.39
2389	South Africa	K	56	68	949	815	233	219	30	2.56
29565	Lesotho	GC	57	66	963	798	242	221	30	2.84
30460	China	C	54	63	921	769	255	206	29	2.42
29582	Lesotho	K	62	69	1049	828	187	150	29	2.44
2872	Egypt	CB	48	60	822	743	181	147	25	3.61
Mean of top 15 accessions			61	71	1025	855	260	227	33	2.68
Range of top 15 accessions			48–70	60–79	822–1175	743–951	181–311	147–285	25–54	2.11–3.61
Mean of Sub-group 2			60	71	1009	862	250	220	24	2.28
Range of Sub-group 2			44–83	54–93	756–1369	685–1108	115–354	106–311	10–54	1.32–3.61
Subgroup 3 (125 accessions)										
1004	India	D	93	84	1539	1007	355	286	49	3.84
23590	Ethiopia	GC	82	72	1358	866	282	220	43	3.04
28141	Yemen	DC	103	75	1688	911	332	249	42	5.18

(Continued)

TABLE 4 | Continued

IS No.	Origin	Race	Days to 50% flowering		Cumulative growing degree days		Plant height (cm)		Grain yield (g)	100-seed weight (g)
			Rainy	Post-rainy	Rainy	Post-rainy	Rainy	Post-rainy		
23891	Yemen	D	99	75	1623	909	394	270	42	5.33
27034	Sudan	D	113	89	1844	1058	365	317	41	3.42
23586	Ethiopia	GC	81	73	1339	874	289	221	40	2.94
31706	Yemen	D	108	74	1782	893	388	308	39	4.69
23514	Ethiopia	C	75	69	1243	826	262	190	38	2.81
21083	Kenya	C	108	77	1776	922	342	297	37	2.58
6421	India	D	109	80	1790	969	371	262	37	2.63
20632	United States	C	127	82	2065	985	390	326	34	3.10
4613	India	D	80	75	1328	896	350	229	32	2.99
19153	Sudan	GC	77	70	1280	836	247	165	30	2.45
9108	Kenya	C	77	71	1281	844	327	237	27	2.66
23579	Ethiopia	GC	77	67	1280	806	323	220	37	2.58
Mean of top 15 accessions			94	75	1548	907	334	253	38	3.35
Range of top 15 accessions			75–127	67–89	1243–2065	806–1058	247–394	165–326	27–49	2.45–5.33
Mean of Sub-group 3			104	74	1704	899	343	247	24	2.63
Range of Sub-group 3			61–148	51–120	1026–2356	658–1485	192–446	120–369	7–49	1.00–5.33
Mean of entire PSTINS			87	73	1433	885	307	236	24	2.99
Range of entire PSTINS			44–148	51–120	756–2356	658–1485	129–446	106–369	7–54	1.00–5.33
Control										
2205	India	DB	76	80	1267	960	334	251	26	2.14
18758	Ethiopia	GC	78	69	1288	833	283	166	29	2.37
33844	India	D	75	77	1239	924	325	262	41	3.60
SE(d)			2.45	1.48	39.88	14.18	16.00	11.16	4.83	0.19
LSD ($P \leq 0.05$)			4.80	2.89	78.17	27.79	31.36	21.87	9.46	0.37
CV (%)			4.18	1.95	4.11	1.53	7.51	4.63	21.45	8.10

Subgroup 2 for the long-day rainy season and subgroup 3 for the short-day post-rainy season.

The CGDD requirements of the Subgroup 2 was on average, 1,009°d in the rainy and 862°d in the post-rainy with a narrow range in both the seasons (756–1,369°d in rainy; 685–1,108°d in post-rainy). The CGDD requirements for the Subgroup 3 landraces was on average 1,704°d in the rainy and 899°d in the post-rainy season with a wide range in the rainy (1,026–2,356°d) and post-rainy (658–1,485°d). Individually the differences in the CGDD requirements in the rainy and post-rainy season of the 125 accessions in the Subgroup 3 were much greater (341–1,475°d) than in the 80 accessions of Subgroup 2 (58–278°d). This clearly established that the PSTINS accessions of the mini core in group 2 and 3 were different in their adaptation in terms of DFL and CGDD requirements.

Agronomic Performance of Mini Core Accessions

The mini core accessions and three control cultivars, two PTINS (IS 205, IS 33844) and one PSTINS (IS 18758), were evaluated for grain yield per plant and 100-seed weight. Vast differences in the grain yield per plant were observed between accessions of the mini core collections in both the post-rainy seasons.

Performance of PTINS Mini Core Sources

The detailed information on the performance of 18 PTINS sources and three control cultivars is given in **Table 2**. The PTINS originated from 11 countries, 6 from India, and 2 each from Ethiopia and the United States and 1 each from eight other countries (**Table 2**). Over both seasons, three PTINS produced a grain yield of 30–33 g per plant, however, none of them was superior to the IS 33844 (**Table 2**). For 100-seed weight also, none of the PTINS (1.43–2.99 g) was superior to IS 33844 (3.6 g) (**Table 2**). Considering DFL in the rainy and post-rainy seasons and yield per plant of IS 29568 appears promising.

Performance of PSTINS Mini Core Sources

The 205 PSTINS accessions originated from 54 countries, six countries contributing 10 or more accessions were Cameroon (12), Ethiopia (10), India (24), Yemen (15), the United States (13) and South Africa (22). Nine accessions in the combined over two post-rainy seasons (39–54 g) performed very well and yielded significantly greater than the PSTINS control IS 18758 (29 g) (**Table 4**). For 100 seed weight, 68 lines had greater 100-seed weight than the IS 18758 (2.4 g).

Subdivision of PSTINS into Subgroups revealed adaptation of Subgroup 2 to the rainy season and Subgroup 3 to the post-rainy season in terms of the DFL and CGDD requirements. This was further confirmed by the performance of these two subgroups in terms of grain yield and 100-seed weight (**Table 4**). Though the overall performance of both the subgroups in terms of mean grain yield per plant was similar (24 g), the number of accessions that produced significantly greater yield than the control IS 18758 were higher (17 in 2010–2011, 13 in 2011–2012, 8 overall) in the

Subgroup 3 compared to subgroup 2 (7 in 2010–2011, 3 in 2011–2012 and 1 overall). In terms of percentage also, since the number of accessions in the two subgroups were different, Subgroup 3 had 13.6% of accessions in 2010–2011, 10.4% in 2011–2012 and 6.4% overall that were significantly superior in yield to IS 18758 compared to 8.5% in 2010–2011, 3.8% in 2011–2012 and only 1.3% in overall in both seasons in Subgroup 2. For 100-seed weight also, a similar pattern was observed with Subgroup 3 showing a greater number of significant accessions (52, 62, and 54) than Subgroup 2 (15, 21, and 14) over the control IS 18758. The two post-rainy seasons had a differential influence on the accessions in both the subgroups. Combined over both seasons, the grain yield varied from 10 to 54 g in Subgroup 2 and 7 to 49 g in Subgroup 3, with a mean value of 24 g for both the Subgroups.

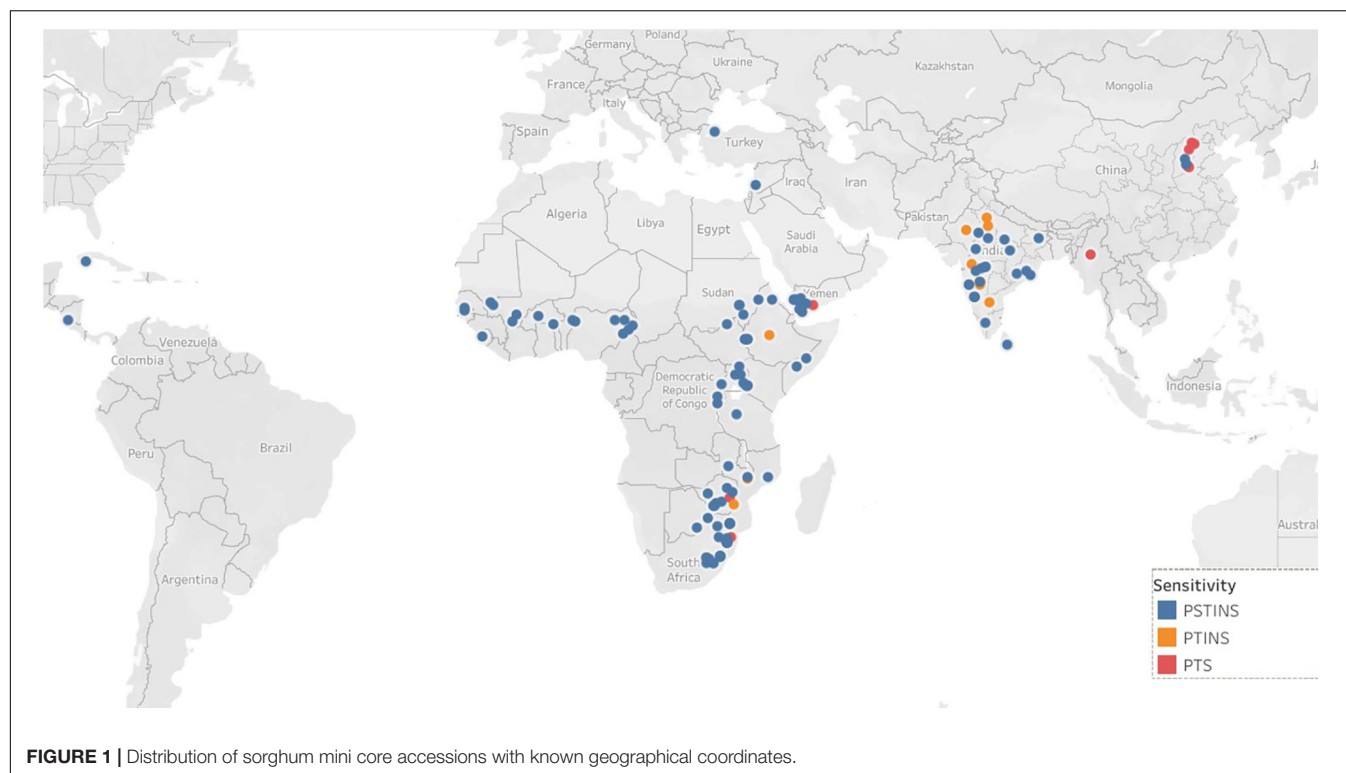
The performance of the top 15 accessions with a desirable combination of DFL, grain yield and 100-seed weight in the Subgroups 2 and 3 are given in **Table 4**. The selected 15 Subgroup 2 sources on average yielded 33 g, 11.4% more than the control IS 18758 (29 g). IS 4698 from India with 54 g was the best line in this group, giving 86.2% more grain yield than IS 18758. Other promising high-yielding sources in this group were IS 22294 from Botswana, IS 29241 and IS 29187 from Swaziland and IS 13971 from South Africa, producing 20.6–24.1% more grain yields than IS 18758 (**Table 4**). In Subgroup 3, the 15 selected accessions yielded on average 38 g, 31% more than IS 18758. Of these, IS 1004 from India, and IS 23586 and IS 23590 from Ethiopia (40–49 g), and IS 31706, IS 23891, and IS 28141 (39–42 g) from Yemen were significantly superior and gave 34–69.0% more yield than IS 18758. IS 31706, IS 23891, and IS 28141 had a good combination of significantly greater yields (39–42 g) and 100 seed weight (4.69–5.33 g) than the control IS 18758 (yield 29 g, 100-seed weight 2.37 g).

Performance of PTS Mini Core Sources

The 19 PTS accessions originated from 9 countries, predominated by China with 8 accessions followed by 2 each from Korea, Zimbabwe and South Africa and 1 each from Cameroon, Yemen, Swaziland, the United States and Myanmar (**Table 3**). The grain yield of 19 PTS ranged from 2 to 29 g in the 2010–2011, 3–33 g in the 2011–2012 and 2–28 g overall in both the seasons. None of the PTS accessions was significantly better than control IS 18758, however, IS 30451 was promising in 2010–2011, 2011–2012 and combined over two post-rainy seasons with grain yields similar to IS 18758 (29 g) (**Table 4**).

Distribution of Sensitive and Insensitive Sources Across Latitudes

Data on the latitude of origin of only 135 mini core accessions, including 11 PTINS, 10 PTS, and 114 PSTINS landraces were available and plotted in **Figure 1**. For the 103 mini core landraces, the hemispheres were decided based on country of origin and the four landraces which originated from the countries spread over both the hemisphere were not included in this analysis. Of the 80 PSTINS of Subgroup 2, latitude data was available only for 41 and in Subgroup 3 on 73 out of 125. Hemisphere wise, 86 accessions (7 PTINS, 71 PSTINS, and 8 PTS) were from North



Hemisphere and 49 (4 PTINS, 43 PSTINS, and 2 PTS) were from South Hemisphere. Among PSTINS distribution of Subgroup 2 and 3 was contrasting, with Subgroup 2 being predominant in the South Hemisphere (29 out of 41) and Subgroup 3 being predominant in the North (59 out of 73). The PTINS sources were distributed from 6 to 30° in the northern hemisphere and from 15° to 30° in the southern hemisphere. Similarly, the PTS sources were distributed from 10 to 40° in the north and from 15 to 30° in the southern hemisphere. Since the number of accessions in the northern and southern hemisphere in the PTINS and PTS was less than 10 in each, it would be meaningless to determine and discuss the patterns. In the case of PSTINS Subgroup 2, all the 29 sources in the southern hemisphere originated from 15 to 30° and all the 59 sources in Subgroup 3 originated from 0 to 30° in the northern hemisphere. The 12 sources of Subgroup 2 were distributed in northern from 11 to 45° and 14 sources in the southern hemisphere originated from 0 to 25°. In both the hemisphere, an increase in latitude results in a decrease in DFL and CGDD in the entire mini core collection. The mean yield among hemispheres was similar in the entire mini core, PTINS, PSTINS, Subgroup 3, and PTS. However, the range was higher in the entire mini core and PSTINS (7–54 g), Subgroup 2 (10–54), and 3 (7–49) in the northern hemisphere than in the southern hemisphere. A similar pattern was observed for the range of 100-seed weight (Table 5).

Correlations Among Quantitative Traits in the Insensitive and Sensitive Groups

Phenotypic correlations were performed among all quantitative traits in the entire mini core, in the PTINS, PSTINS and

PTS groups and two subgroups of the PSTINS separately (Table 6). The DFL in the rainy and post-rainy seasons were highly correlated with plant height in the rainy and post-rainy seasons in the entire core, PTINS, entire PSTINS, PSTINS Subgroup 2 and Subgroup 3 ($p < 0.001$ –0.05), but it was non-significant in the PTS group. Contrastingly, grain yield in the post-rainy season was highly correlated with DFL, CGDD and plant height in the rainy and post-rainy seasons ($r = 0.302$ –0.413, $p < 0.01$ –0.001) only in the PSTINS Subgroup 2 and non-significant in other groups except entire mini core and PSTINS entire group. In the latter two groups, the entire mini core and PSTINS group the magnitude of correlations was low at 0.187 and 0.190, respectively. Further, in the PTS group the highly significant correlations were observed only between 100-seed weight and grain yield in post-rainy, plant height between rainy and post-rainy, and DFL in the rainy and post-rainy seasons, all other correlations were non-significant (Table 6). This demonstrated that the strength of correlations depended upon the population used under study. The more homogenous the group, the better were the estimate of correlation. Further, the proportion of variance in one trait that can be attributed to its linear relationship with a second trait is indicated by the square of the correlation coefficient (Snedecor and Cochran, 1980). Estimates of this value greater than 0.71 or lower than -0.71 have been suggested as meaningful correlations (Skinner et al., 1999). In our study, we found such high correlation coefficients between plant height in the rainy and post-rainy seasons in all six groups (Table 6).

DISCUSSION

The response of crops to the prevailing photoperiod and temperature conditions determines their adaptation to the wide range of different environments in which it is cultivated worldwide. Sorghum is a short-day species in which flowering initiation occurs only during the short-day condition (Oliveira et al., 2019). The majority of accessions conserved at the ICRISAT genebank do not flower or take much more time when they are being grown in the long-day rainy season (June–October) in India, while most accessions flower in the short-day post-rainy (October–February) season. In this study, the

sorghum mini core accessions were characterized for traits related to photoperiod and temperature sensitivity, revealed a significant genotypic variation for DFL, CGDD, plant height, grain yield and 100-seed weight, and the interactions between genotype and season (year) were significant. This indicated that the sorghum mini core collection had adequate genotypic variation and the performance of genotypes differed in the two seasons.

Considering the DFL and CGDD, the mini core accessions were categorized into photoperiod and temperature insensitive (PTINS) (18 accessions), photoperiod sensitive and temperature insensitive (PSTINS) (205 accessions), and photoperiod and

TABLE 5 | Hemisphere wise, mean and range of days to 50% flowering, cumulative growing degree days (CGDD), plant height, grain yield and 100-seed weight of the entire mini core and sensitivity response groups.

Trait/group	Northern hemisphere				Southern hemisphere			
	Mean		Range		Mean		Range	
	Rainy	Post-rainy	Rainy	Post-rainy	Rainy	Post-rainy	Rainy	Post-rainy
Entire mini core (160 accessions northern hemisphere; 77 accessions southern hemisphere)								
Days to 50% flowering	87	71	44–148	51–120	73	76	49–143	58–100
CGDD	1430	863	756–2356	658–1485	1225	913	839–2289	728–1202
Plant height (cm)	315	238	129–446	120–369	270	229	115–399	106–313
Grain yield g plant ⁻¹		23		7–54		24		2–37
100-Seed weight (g)		2.53		1.00–5.33		2.2		0.48–3.42
PTINS (11 accessions northern hemisphere; 7 accessions southern hemisphere)								
Days to 50% flowering	69	69	57–84	56–87	69	69	61–82	58–78
CGDD	1159	840	971–1386	711–1041	1159	832	1025–1350	728–941
Plant height (cm)	283	224	185–335	129–331	270	215	189–335	137–260
Grain yield g plant ⁻¹		22		14–33		21		10–29
100-Seed weight (g)		2.20		1.48–2.89		2.14		1.43–2.99
PSTINS (125 accessions northern hemisphere; 65 accessions southern hemisphere)								
Days to 50% flowering	92	72	44–148	51–120	75	76	54–143	58–100
CGDD	1514	871	756–2356	658–1485	1246	913	910–2289	728–1202
Plant height (cm)	324	239	129–446	120–369	271	230	115–399	106–313
Grain yield g plant ⁻¹		24		7–54		25		9–37
100-Seed weight (g)		2.61		1.00–5.33		2.30		1.42–3.42
PSTINS Subgroup 2 (36 accessions northern hemisphere; 44 accessions southern hemisphere)								
Days to 50% flowering	58	69	44–83	54–93	61	74	54–78	62–92
CGDD	974	830	756–1369	685–1108	1037	888	910–1298	766–1096
Plant height (cm)	259	220	129–354	132–311	242	220	115–319	106–295
Grain yield g plant ⁻¹		22		10–54		27		15–36
100-Seed weight (g)		2.17		1.32–3.61		2.38		1.42–3.42
PSTINS Subgroup 3 (99 accessions northern hemisphere; 21 accessions southern hemisphere)								
Days to 50% flowering	104	73	61–148	51–120	103	80	66–143	58–100
CGDD	1710	887	1026–2356	658–1485	1683	964	1112–2289	728–1202
Plant height (cm)	347	246	192–446	120–369	331	250	211–399	166–313
Grain yield g plant ⁻¹		24		7–49		22		9–37
100-Seed weight (g)		2.77		1.00–5.33		2.13		1.42–3.21
PTS (14 accessions northern hemisphere; 5 accessions southern hemisphere)								
Days to 50% flowering	49	66	46–62	60–88	62	87	49–66	73–94
CGDD	832	804	785–1052	738–1054	1039	1037	839–1110	880–1121
Plant height (cm)	258	240	214–293	202–287	254	235	189–298	156–278
Grain yield g plant ⁻¹		20		14–28		18		2–25
100-Seed weight		1.99		1.33–2.43		1.70		0.48–2.50

temperature-sensitive (PTS) (19 accessions) and identified best performing accessions in each of the group for their utilization in crop improvement. The 19 PTS sources identified in this study flowered 14–28 days earlier in the long-day rainy season than in the short-day post-rainy season, but with similar CGDD requirement in the two seasons. Among PSTINS accessions, 80 accessions were in Subgroup 2 that flowered 5–22 days earlier

in the rainy season than in the post-rainy season, but with significantly higher CGDD requirement during the rainy season. These two groups (PTS and PSTINS subgroup 2) may have different mechanisms to trigger flowering than the 18 PTINS sources and the 125 PSTINS Subgroup 3 sources. The important question then is “Are there landraces in the global collections which are adapted to long-days such as PTS and PSTINS

TABLE 6 | Correlations among different traits in the entire mini core collection and sensitivity response groups.

Trait		DF-R [†]	DF-PR	CGDD-R	CGDD-PR	PLHT-R	PLHT-PR	GY-PR
DF-PR	Entire (242)	0.485***						
	PTINS (18)	0.927***						
	PSTINS (205)	0.492***						
	Sub-group 2 (80)	0.860***						
	Sub-group 3 (125)	0.591***						
	PTS (19)	0.976***						
CGDD-R	Entire (242)	1.000***	0.484***					
	PTINS (18)	0.999***	0.927***					
	PSTINS (205)	1.000***	0.491***					
	Sub-group 2 (80)	0.994***	0.861***					
	Sub-group 3 (125)	1.000***	0.593***					
	PTS (19)	1.000***	0.979***					
CGDD-PR	Entire (242)	0.489***	0.997***	0.488***				
	PTINS (18)	0.934***	0.996***	0.934***				
	PSTINS (205)	0.495***	0.997***	0.494***				
	Sub-group 2 (80)	0.850***	0.997***	0.849***				
	Sub-group 3 (125)	0.583***	0.997***	0.585***				
	PTS (19)	0.976***	1.000***	0.979***				
PLHT-R	Entire (242)	0.744***	0.326***	0.745***	0.327***			
	PTINS (18)	0.510*	0.541*	0.515*	0.495*			
	PSTINS (205)	0.745***	0.320***	0.747***	0.324***			
	Sub-group 2 (80)	0.427***	0.415***	0.410***	0.416***			
	Sub-group 3 (125)	0.520***	0.256**	0.526***	0.252**			
	PTS (19)	0.119	0.132	0.123	0.139			
PLHT-PR	Entire (242)	0.473***	0.570***	0.472***	0.559***	0.724***		
	PTINS (18)	0.705***	0.817***	0.710***	0.794***	0.838***		
	PSTINS (205)	0.517***	0.582***	0.517***	0.573***	0.735***		
	Sub-group 2 (80)	0.352***	0.487***	0.344**	0.489***	0.873***		
	Sub-group 3 (125)	0.582***	0.612***	0.586***	0.596***	0.731***		
	PTS (19)	0.09	0.155	0.098	0.165	0.896***		
GY-PR	Entire (242)	0.088	0.119	0.089	0.09	0.109	0.187**	
	PTINS (18)	0.34	0.466	0.336	0.434	0.145	0.242	
	PSTINS (205)	0.02	0.092	0.021	0.061	0.067	0.190**	
	Sub-group 2 (80)	0.379***	0.401***	0.394***	0.387***	0.302**	0.413***	
	Sub-group 3 (125)	0.027	−0.015	0.027	−0.047	0.029	0.104	
	PTS (19)	0.182	0.047	0.174	0.046	0.148	0.023	
HSW-PR	Entire (242)	0.239***	−0.203***	0.239***	−0.204***	0.229***	−0.012	0.492***
	PTINS (18)	−0.118	0.014	−0.118	−0.013	−0.229	−0.153	0.747***
	PSTINS (205)	0.180**	−0.248***	0.180**	−0.247***	0.207**	−0.003	0.454***
	Sub-group 2 (80)	−0.049	−0.077	−0.04	−0.065	−0.106	−0.171	0.237*
	Sub-group 3 (125)	0.02	−0.350***	0.016	−0.353***	0.151	−0.034	0.553***
	PTS (19)	0.009	−0.16	−0.006	−0.161	−0.058	−0.267	0.698***

[†]DF-R, days to 50% flowering in rainy; DF-PR, days to 50% flowering in post-rainy; CGDD-R, cumulative growing degree days in rainy; CGDD-PR, cumulative growing degree days in post-rainy; PLHT-R, Plant height (cm) in rainy; PLHT-PR, Plant height (cm) in post-rainy; GY-PR, Grain yield (g plant^{−1}) in post-rainy; HSW-PR, 100-seed weight in post-rainy; *, **, *** significant at $P \leq 0.05, 0.01, 0.001$, respectively.

Subgroup 2 sources in the present study?” A critical analysis of flowering under controlled temperature and photoperiod conditions may provide the answer to whether flowering in some sorghum accessions is triggered by long-day or a combination of long-day \times temperature interaction. In most photoperiod and temperature studies under controlled conditions, only a few cultivars have been used to model the flowering, which agrees with the known hypothesis that sorghum is a short-day plant, which may not be true when diverse landraces representing the diversity of global collection are used as shown in the present study with mini core collection.

In the entire collection of sorghum (Upadhyaya et al., 2018), and also in the mini core collection in this study, the mean and range of DGL and CGDD requirements for flowering decreased with the increase in latitude in both the hemisphere. The magnitude of the relationship between DFL and latitude differed between groups (PTINS, PSTINS, and PTS groups and two subgroups of the PSTINS separately) indicating that it varied with the constitution of the group. The differences in the relationship between northern and southern hemispheres in the mini core could be attributed to the differences in hemispheres in terms of the proportion of land and water and associated climatic parameters (Upadhyaya et al., 2018). The northern hemisphere has four times land portion (80%) than the ocean (20%), and land surfaces conduct the heat poorly, therefore, most of the absorbed radiation heats the land surface and associated vegetation in the northern hemisphere. In addition high levels of man-made pollution also contribute to the higher temperature in the northern hemisphere. The southern hemisphere has one-and-half times ocean portion (60%) than land (40%), and water is a good conductor of heat, and most of the radiation absorbed is stored in the deeper waters in the southern hemisphere.

Plant breeders in their quest to develop high-yielding cultivars with a broad adaptation need to work on several traits simultaneously. It is therefore important to have sources that facilitate developing such cultivars. In our efforts toward identifying such sources, this study was undertaken to identify photoperiod and temperature insensitive sources in the sorghum mini core collection. In earlier studies, sorghum mini core accessions have been evaluated for biotic (Sharma et al., 2010; Radwan et al., 2011; Seifers et al., 2012; Borphukan, 2014) and abiotic (Kapanigowda et al., 2013; Upadhyaya et al., 2016c, 2017) stresses, and for grain nutritional (Upadhyaya et al., 2016a) and bioenergy traits (Wang et al., 2011; Upadhyaya et al., 2014a). Several photoperiod- and/or temperature-insensitive sources identified in this study has also been reported as sources of other important traits for the breeding program. Of the 18 PTINS sources, 15 have been reported as sources for one or more traits. IS 5094 has been reported tolerant to drought based on the drought tolerance index (Upadhyaya et al., 2017), resistant to downy mildew (Sharma et al., 2010), charcoal rot, stem borer and shoofly. Similarly, IS 473, another PTINS was found resistant to grain mold, Anthracnose, leaf blight and rust (Sharma et al., 2010, 2012). These and other PTINS sources with resistance could be good sources to breed sorghum cultivars with stress resistance and insensitivity to photoperiod and temperature to impart wide adaptation. A large number of PSTINS sources, 65

out of 80 in the Subgroup 2 and 89 out of 125 in Subgroup 3 had shown desirability to one or more traits. Four Subgroup 2 accessions, IS 602 (Zinc, germination at low temperature, grain mold, and charcoal rot), IS 1233 (Iron, Zinc, germination at low temperature, and grain mold), IS 4515 (drought, stem borer, shoot fly, and aphids), and IS 4698 (dual purpose, stem borer, shoot fly, and aphids) had desirability for four traits each. Similarly, three accessions, IS 22294 and IS 30092 for downy mildew, grain mold and charcoal rot and IS 12945 for grain mold, leaf blight and charcoal rot. These sources could be useful in developing resistant sorghum cultivars suitable for rainy season. The 89 Subgroup 3 sources had shown desirability for large number of traits. One line, IS 23684 for six traits (protein, Brix, anthracnose, leaf blight, rust, and Aphids), five accessions for four traits: IS 7305 for Saccharification yield, seedling vigor at low temperature, germination at low temperature, downy mildew; IS 23521 for Anthracnose, leaf blight, rust and Aphids; IS 23992 for iron, downy mildew, grain mold, and stem borer; IS 24939 for Brix, Anthracnose, leaf blight, and Aphids and IS 31557 for downy mildew, leaf blight, charcoal rot, and Aphids. These and other sources in this group would be a valuable resource for breeders to develop multipurpose sorghum cultivars with adaptation to the post-rainy season. Among the 19 PTS sources, 15 have been identified as desirable for 1–5 traits. Two accessions for five traits, IS 1212 for Iron, Zinc, seedling vigor at low temperature, downy mildew and grain mold and IS 24503 for Iron, Zinc, leaf blight, rust, and charcoal rot. IS 29314 (germination at low temperature, downy mildew, and aphids), IS 30466 (Zinc, downy mildew, and charcoal rot) and IS 30507 (Iron, Zinc, and charcoal rot) were good for three traits. These and other sources in this group could be useful sources to the breeders to develop cultivars suitable for the areas where both photoperiod and temperatures are important in sorghum cultivation and can be exploited to enhance production.

The use of novel variability is the key to continued progress in crop improvement and for widening the genetic base of the newly bred cultivars (Upadhyaya et al., 2009). The information on the differential response of sorghum accessions in response to photoperiod and temperature sensitivity could be useful in the sorghum improvement program. The PTINS, PTS and PSTINS sources identified in this study could help breeders to use appropriate sources in their program. The photoperiod and temperature insensitive accessions can be utilized to developed cultivars with broader adaptation, while the highly photoperiod sensitive tall accessions can be utilized for biomass and forage improvement. Further, the information on the desirability of the identified accessions for other important traits will be useful to determine the accessions based on needs for extensive multilocation evaluation in the targeted environments to identify stable sources for breeding programs. The source accessions identified in the present study could be utilized in genetic and genomic investigations to understand the mechanism underlying photoperiod and temperature response and mapping the genes. Sorghum researchers can obtain seed samples of these accessions from the ICRISAT genebank for research purposes through a Standard Material Transfer Agreement (SMTA).

DATA AVAILABILITY STATEMENT

The original contributions generated for this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

HU and MV: conceptualization. MV: data analysis and curation. HU: writing – original draft. HU, MV, and VA: writing – review and editing. All authors contributed to the article and approved the submitted version.

REFERENCES

- Assefa, Y., Staggenborg, S. A., and Prasad, V. P. V. (2010). Grain sorghum water requirement and responses to drought stress: a review. *Crop Manag.* 9, 1–12. doi: 10.1094/cm-2010-1109-01-rv
- Borphan, B. (2014). *Evaluation of Minicore Germplasm of Rabi Sorghum for Charcoal Rot Resistance and Yield Component Traits, Expression Analysis of Selected R-Genes During Charcoal Rot Disease Incidence*. M.Sc. thesis. Dharwad: College of Agriculture, Dharwad.
- Dalrymple, D. G. (1986). *Development and Spread Of High-Yielding Wheat Varieties In Developing Countries*, 7th Edn. Washington, DC: U.S. Agency for International Development, doi: 10.1016/0304-1131(74)90052-6
- Doggett, H. (1988). *Sorghum*, 2nd Edn. Harlow, UK: Longman Scientific & Technical.
- Gerik, T., Bean, B., and Vanderlip, R. (2003). *Sorghum Growth and Development*. Publ. B-6137. *AgriLife Communications and Marketing*, College Station, TX: A&M System.
- Jordan, D. R., Tao, Y. Z., Godwin, I. D., Henzell, R. G., Cooper, M., and McIntyre, C. L. (1998). Loss of genetic diversity associated with selection for resistance to sorghum midge in Australian sorghum. *Euphytica* 102, 1–7. doi: 10.1023/A:1018311908636
- Jordan, D. R., Tao, Y., Godwin, I. D., Henzell, R. G., Cooper, M., and McIntyre, C. L. (2003). Prediction of hybrid performance in grain sorghum using RFLP markers. *Theor. Appl. Genet.* 106, 559–567. doi: 10.1007/s00122-002-1144-5
- Kapanigowda, M. H., Perumal, R., Djanaguiraman, M., Aiken, R. M., Tesso, T., Prasad, P. V., et al. (2013). Genotypic variation in sorghum [*Sorghum bicolor* (L.) Moench] exotic germplasm collections for drought and disease tolerance. *Springerplus* 2:650. doi: 10.1186/2193-1801-2-650
- Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4:2320. doi: 10.1038/ncomms3320
- Muraya, M. M., de Villiers, S., Parzies, H. K., Mutegi, E., Sagnard, F., Kanyenji, B. M., et al. (2011). Genetic structure and diversity of wild sorghum populations (*Sorghum* spp.) from different eco-geographical regions of Kenya. *Theor. Appl. Genet.* 123, 571–583. doi: 10.1007/s00122-011-1608-6
- Murray, S. C., Rooney, W. L., Hamblin, M. T., Mitchell, S. E., and Kresovich, S. (2009). Sweet sorghum genetic diversity and association mapping for Brix and height. *Plant Genome* 2, 48–62. doi: 10.3835/plantgenome2008.10.0011
- Oliveira, I. C. M., Marçal, T. D. S., Bernardino, K. D. C., Ribeiro, P. C. D. O., Parrella, R. A. D. C., Carneiro, P. C. S., et al. (2019). Combining ability of biomass sorghum lines for agroindustrial characters and multitrait selection of photosensitive hybrids for energy cogeneration. *Crop Sci.* 59, 1554–1566. doi: 10.2135/cropsci2018.11.0693
- Patterson, H. D., and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554. doi: 10.1093/biomet/58.3.545
- Radwan, G. L., Perumal, R., Isakeit, T., Magill, C. W., Prom, L. K., and Little, C. R. (2011). Screening exotic sorghum germplasm, hybrids, and elite lines for resistance to a new virulent pathotype (P6) of *Peronosclerospora sorghi*

FUNDING

We gratefully acknowledge support from the CGIAR Genebank Platform coordinated by the Crop Trust for supporting Genebank activities.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.571243/full#supplementary-material>

- causing downy mildew. *Plant Heal. Prog* 12:1. doi: 10.1094/PHP-2011-0323-01-RS
- Seifers, D. L., Perumal, R., and Little, C. R. (2012). New sources of resistance in sorghum (*Sorghum bicolor*) germplasm are effective against a diverse array of Potyvirus spp. *Plant Dis.* 96, 1775–1779. doi: 10.1094/PDIS-03-12-0224-RE
- Sharma, H. C., Taneja, S. L., Rao, N. K., and Rao, K. E. P. (2003). *Evaluation of Sorghum Germplasm for Resistance to Insect Pests*. Information Bulletin no. 63. Patancheru 502 324, Andhra Pradesh, India. Patancheru: International Crops Research Institute for the Semi-Arid Tropics, 184.
- Sharma, R., Rao, V. P., Upadhyaya, H. D., Reddy, V. G., and Thakur, R. P. (2010). Resistance to grain mold and downy mildew in a mini-core collection of sorghum germplasm. *Plant Dis.* 94, 439–444. doi: 10.1094/PDIS-94-4-0439
- Sharma, R., Upadhyaya, H. D., Manjunatha, S. V., Rao, V. P., and Thakur, R. P. (2012). Resistance to foliar diseases in a mini-core collection of sorghum germplasm. *Plant Dis.* 96, 1629–1633. doi: 10.1094/pdis-10-11-0875-re
- Skinner, D. Z., Bauchan, G. R., Auricht, G., and Hughes, S. (1999). “Developing core collections for a large annual Medicago collection,” in *Core Collection For Today And Tomorrow*, eds R. C. Johnson and T. Hodgkin (Rome: Int. Plant Genetic Resources Inst), 61–67.
- Smith, S., Primomo, V., Monk, R., Nelson, B., Jons, E., and Porter, K. (2010). Genetic diversity of widely used U.S. sorghum hybrids 1980–2008. *Crop Sci.* 50, 1664–1673. doi: 10.2135/cropsci2009.10.0619
- Snedecor, G. W., and Cochran, W. G. (1980). *Statistical Methods*, 7th Editio Edn. Iowa: Iowa State University Press.
- Upadhyaya, D., and Ortiz, R. (2001). A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor. Appl. Genet.* 102, 1292–1298. doi: 10.1007/s00122-001-0556-y
- Upadhyaya, H. D., Dronavalli, N., Dwivedi, S. L., Kashiwagi, J., Krishnamurthy, L., Pande, S., et al. (2013). Mini core collection as a resource to identify new sources of variation. *Crop Sci.* 53, 2506–2517. doi: 10.2135/cropsci2013.04.0259
- Upadhyaya, H. D., Dwivedi, S. L., Ramu, P., Singh, S. K., and Singh, S. (2014a). Genetic variability and effect of postflowering drought on stalk sugar content in sorghum mini core collection. *Crop Sci.* 54, 2120–2130. doi: 10.2135/cropsci2014.01.0040
- Upadhyaya, H. D., Dwivedi, S. L., Sharma, S., Lalitha, N., Singh, S., Varshney, R. K., et al. (2014b). Enhancement of the use and impact of germplasm in crop improvement. *Plant Genet. Resour.* 12, S155–S159. doi: 10.1017/S1479262114000458
- Upadhyaya, H. D., Dwivedi, S. L., Singh, S., Sahrawat, K. L., and Singh, S. K. (2016a). Genetic variation and postflowering drought effects on seed iron and zinc in ICRISAT sorghum mini core collection. *Crop Sci.* 56, 374–383. doi: 10.2135/cropsci2015.05.0308
- Upadhyaya, H. D., Dwivedi, S. L., Vadez, V., Hamidou, F., Singh, S., Varshney, R. K., et al. (2014c). Multiple resistant and nutritionally dense germplasm identified from mini core collection in peanut. *Crop Sci.* 54, 679–693. doi: 10.2135/cropsci2013.07.0493
- Upadhyaya, H. D., Dwivedi, S. L., Vetriventhan, M., Krishnamurthy, L., and Singh, S. K. (2017). Post-flowering drought tolerance using managed stress trials, adjustment to flowering, and mini core collection in sorghum. *Crop Sci.* 57, 1–12. doi: 10.2135/cropsci2016.04.0280

- Upadhyaya, H. D., Pundir, R. P. S., Dwivedi, S. L., Gowda, C. L. L., Reddy, V. G., and Singh, S. (2009). Developing a mini core collection of sorghum for diversified utilization of germplasm. *Crop Sci.* 49, 1769–1780. doi: 10.2135/cropsci2009.01.0014
- Upadhyaya, H. D., Reddy, K. N., Vetriventhan, M., and Ahmed, M. I. (2018). Latitudinal adaptation of flowering response to photoperiod and temperature in the world collection of sorghum landraces. *Crop Sci.* 58, 1265–1276. doi: 10.2135/cropsci2017.05.0298
- Upadhyaya, H. D., Sharma, S., Dwivedi, S. L., and Singh, S. K. (2014d). “Sorghum genetic resources: conservation and diversity assessment for enhanced utilization in sorghum improvement,” in *Genetics, Genomics and Breeding of Sorghum*, eds Y.-H. Wang, H. D. Upadhyaya, and C. Kole (CRC Press Inc), 28–55.
- Upadhyaya, H. D., Vetriventhan, M., and Deshpande, S. (2016b). “Sorghum germplasm resources characterization and trait mapping,” in *The Sorghum Genome*, eds S. Rakshit and Y.-H. Wang (Springer International Publishing), doi: 10.1007/978-3-319-47789-3_4
- Upadhyaya, H. D., Vetriventhan, M., Asiri, A. M., and Azevedo, V. C. R. (2019). Multi-trait diverse germplasm sources from mini core collection for sorghum improvement. *Agriculture* 9:121. doi: 10.3390/agriculture9060121
- Upadhyaya, H. D., Wang, Y.-H., Dintyala, S. V., Dwivedi, S. L., Prasad, P. V. V., Burrell, A. M., et al. (2016c). Association mapping of germinability and seedling vigor in sorghum under controlled low temperature conditions. *Genome* 59, 137–145. doi: 10.1139/gen-2015-0122
- Vellve, R. (1992). *Saving the Seeds: Genetic Diversity And European Agriculture*. London, U.K: Earthscan Publication.
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Am. Math. Soc.* 54, 426–482. doi: 10.1090/s0002-9947-1943-0012401-3
- Wang, Y.-H., Poudel, D. D., and Hasenstein, K. H. (2011). Identification of SSR markers associated with saccharification yield using pool-based genome-wide association mapping in sorghum. *Genome* 54, 883–889. doi: 10.1007/s11032-011-9617-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Upadhyaya, Vetriventhan and Azevedo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Harnessing Sorghum Landraces to Breed High-Yielding, Grain Mold-Tolerant Cultivars With High Protein for Drought-Prone Environments

Mallela Venkata Nagesh Kumar^{1*}, Vittal Ramya¹, Mahalingam Govindaraj², Chanda Venkata Sameer Kumar¹, Setaboyine Maheshwaramma¹, Seshu Gokenpally¹, Mathyam Prabhakar³, Hariprasanna Krishna⁴, Mulinti Sridhar¹, Maparla Venkata Ramana¹, Kodari Avil Kumar¹ and Rumandla Jagadeeshwar¹

¹ Regional Agricultural Research Station, Palem - Professor Jayashankar Telangana State Agricultural University, Hyderabad, India, ² Sorghum Improvement Asia Program - International Crop Research Institute for Semi Arid Tropics, Patancheru, India, ³ Indian Council of Agricultural Research—Central Research Institute for Dryland Agriculture, Hyderabad, India, ⁴ Indian Council of Agricultural Research—Indian Institute of Millets Research, Hyderabad, India

OPEN ACCESS

Edited by:

Thomas Miedaner,
University of Hohenheim, Germany

Reviewed by:

Surinder Banga,
Punjab Agricultural University, India
Arindam Ghatak,
University of Vienna, Austria

*Correspondence:

Mallela Venkata Nagesh Kumar
mvnkumar.gpbr@gmail.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 28 January 2021

Accepted: 31 May 2021

Published: 30 June 2021

Citation:

Nagesh Kumar MV, Ramya V, Govindaraj M, Sameer Kumar CV, Maheshwaramma S, Gokenpally S, Prabhakar M, Krishna H, Sridhar M, Venkata Ramana M, Avil Kumar K and Jagadeeshwar R (2021) Harnessing Sorghum Landraces to Breed High-Yielding, Grain Mold-Tolerant Cultivars With High Protein for Drought-Prone Environments. *Front. Plant Sci.* 12:659874. doi: 10.3389/fpls.2021.659874

Intermittent drought and an incidence of grain mold disease are the two major constraints affecting sorghum production and productivity. The study aimed at developing drought-tolerant sorghum varieties possessing a high protein content and tolerance to grain mold with stable performance using additive main effects and multiplicative interaction (AMMI) and genotype and genotype \times environment interaction (GGE) biplot methods. Systematic hybridization among the 11 superior landraces resulted in subsequent pedigree-based breeding and selection from 2010 to 2015 evolved 19 promising varieties of grains such as white, yellow, and brown pericarp grains. These grain varieties were evaluated for their adaptability and stability for yield in 13 rainfed environments and for possessing tolerance to grain mold in three hot spot environments. A variety of yellow pericarp sorghum PYPS 2 (3,698 kg/ha; 14.52% protein; 10.70 mg/100g Fe) possessing tolerance to grain mold was identified as a stable variety by using both AMMI and GGE analyses. Four mega-environments were identified for grain yield and fodder yield. Sorghum varieties PYPS 2, PYPS 4, PYPS 8, and PYPS 11 were highly stable in E2 with a low grain mold incidence. Besides meeting the nutritional demand of smallholder farmers under dryland conditions, these varieties are suitable for enhancing sorghum productivity under the present climate change scenario.

Keywords: sorghum, landraces, drought tolerance, grain mold, $G \times E$ interaction, AMMI, GGE biplot

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is a widely adaptable cereal crop cultivated in tropical, subtropical, and temperate regions of the world. It is the fifth most important cereal crop next to wheat, rice, maize, and barley, and it is a staple food for millions of people in the semiarid regions of Asia and Africa (Mundia et al., 2019). Sorghum grew globally in over 40.07 Mha with a production of 57.89 million tons and productivity of 1,444 kg/ha (FAOSTAT, 2019). In India, it is the third most

important food grain next to rice and wheat cultivated in an area of 4.09 Mha (FAOSTAT, 2019) in the rainy (*Kharif*) season as a rainfed crop and in the post-rainy (*Rabi*) season under residual soil moisture. More than 90% of the sorghum area and 85% of the production are concentrated in the warm semiarid tracts of central and south India (Davis et al., 2019). The productivity of sorghum in India is still considered to be low at 849 kg/ha compared to the average global productivity of 1,444 kg/ha (FAOSTAT, 2019). This can be attributed to poor soils (marginal lands), unreliable rainfall, incidence of insect pests and diseases, and poor crop input management. The water deficit is significantly increased due to irregularities in rainfall distribution exacerbated by climate change (Eggen et al., 2019; Ocheing et al., 2020). While the rainy season is predominated by the sorghum hybrids, the post-rainy season is dominated by open-pollinated varieties that contribute to low sorghum productivity (Patil et al., 2014). This indicates that a greater emphasis is required for separating and strengthening the focus of rainy and post-rainy breeding on the cultivars' genetic enhancement and their adaptability.

Sorghum has an inherent ability to adapt to a harsh climate. The crop can grow well in dry conditions and can also tolerate water logging, thus making it ideal for cultivation in the arid and semiarid regions of the world (Hadebe et al., 2017). These factors change in relation to climate change, which is predicted to make sorghum production riskier, especially under rainfed agriculture and more so for smallholder farmers (McCarthy and Vlek, 2012). A feasible approach of modifying management practices through a deliberate choice between an improved sorghum variety and local landraces accompanied by an appropriate time of sowing will enhance the adaptive capacity of many resource-poor sorghum farmers and ultimately increase sorghum production ensuring food and livelihood security (Akinseye et al., 2020).

Grain mold is one of the most important diseases in sorghum, which is caused by a complex of fungal species. The genera *Fusarium*, *Curvularia*, and *Alternaria* are mainly responsible for 80–90% of the infection in India (Das et al., 2020). Species of *Bipolaris*, *Phoma*, *Drechslera*, *Exserohilum*, *Aspergillus*, *Cladosporium*, *Penicillium*, *Olpitrichum*, *Gonatobotrytis*, and *Aspergillus* are also detected sporadically in low frequency (Das et al., 2020). The disease can cause yield losses ranging from 30 to 100% depending on the cultivars and weather conditions (Kalaria et al., 2020). Losses in seed weight, grain density, germination, and seed viability due to grain mold lead to a significant decline in seed quality parameters in the rainy season (Nida et al., 2019). Toxins and secondary metabolites produced by the fungi on infected grains render sorghum unfit for human consumption and for cattle and poultry feed (Das et al., 2020). Host-plant resistance is the most cost-effective, efficient, and eco-friendly management practice (Mofokeng et al., 2017). In sorghum, properties such as panicle compactness, glume cover, glume pigmentation, grain hardness, polyphenols (tannins), flavonoids (flavan-4-ols), and antifungal proteins (chitinases, glucanases, sormatin, PR-10, and RIPs) confer a resistance to grain mold disease. There is a need to avoid dependency on a few sources of grain mold-resistant genes and alleles that are currently available, and in this context, crop wild relatives and landraces

offer a tremendous scope acting as reservoirs of useful genes for sorghum improvement (Brar and Khush, 2018; Kyrtatzis et al., 2019). With an increase in the effect of climate change, there is a need to collect, screen, and identify novel sorghum germplasm harboring the grain mold-resistant trait that can be harnessed for adaptation to rainfed agroecologies of India.

Sorghum is a significant source of dietary energy, protein, and micronutrients for the vast majority of the population in sub-Saharan Africa and India (Awika, 2017). It is a good source of phytochemicals including phenolic acid, flavonoids, anthocyanins, phytosterols, plicosanols, tannins, and carotenoids, which make the grain suitable for developing the functional food and nutraceuticals (Balcerek et al., 2020). Additionally, high antioxidant levels in pigmented and tannin sorghum varieties offer many health benefits including slow digestibility, cholesterol-lowering, antioxidant, anti-inflammatory, and anticarcinogenic properties (Abdelhalim et al., 2021). The presence of tannins in the testa, which is a layer beneath the pericarp, improves a resistance to grain mold in sorghum (Cuevas et al., 2018). However, red-pigmented testa and high tannin content are less desired in India where red grain sorghum varieties are rarely used for human consumption. Yellow pericarp sorghum is rich in flavanones and has slightly higher total phenolic contents than white sorghum (Dykes et al., 2011). There is a great demand for sorghum with a yellow pericarp owing to high nutritional and good flatbread making and keeping qualities (Jaisimha, 2019). Biofortification of sorghum through genetic approaches and an increased intake of nutrition-rich sorghum grains can help in improving the nutritional security in the developing world (Kumar et al., 2015).

Keeping in view of a narrow genetic diversity for grain minerals in modern sorghum cultivars, the identification and utilization of valuable alleles in wild ancestors of crop plants are considered as a sustainable approach for enhancing sorghum nutrition (Mofokeng et al., 2018; Abdelhalim et al., 2019). India is considered as the secondary center of sorghum diversity next to East Africa (Ananda et al., 2020). High levels of within- and between-population variability among the sorghum landrace collections indicate a high germplasm diversity and a traits-based genetic novelty, which contribute to sorghum yield and adaptation improvement (Ghebru et al., 2002). However, adaptability and stability in yield are often challenged by the presence of a genotype \times environment ($G \times E$) interaction, which is a number one factor responsible for changing the genotype performance in different environments. Hence, it is important that multi-environment trials are conducted periodically to investigate the $G \times E$ interaction for selecting stable genotypes for yield and other important traits. The additive main effects and multiplicative interaction (AMMI) analysis proposed by Gauch (2013) and the genotype and $G \times E$ interaction (GGE) biplot model developed by Yan et al. (2000) are powerful tools used by plant breeders, geneticists, and agronomists for the identification of genotypes with high yield and wide adaptability.

These methods were also used to identify the landraces with yield stability and adaptability in sorghum (Admas and Tesfaye, 2017), chickpea (Pouresmael et al., 2018), wheat

(Bavandpori et al., 2018), common bean (Philipo et al., 2020), etc. for further use in breeding programs for the development of new varieties. In sorghum, the high genetic and phenotypic diversity were reported from the landraces collected from India (Elangovan et al., 2009, 2012; Vara Prasad and Sridhar, 2019), Ethiopia (Adugna, 2014; Amelework et al., 2015; Derese et al., 2018; Wondimu et al., 2020), Eritrea (Tesfamichael et al., 2014), and Sudan (Abdelhalim et al., 2021). These landraces are the indispensable sources of genetic variation that can be utilized by plant breeders in the development of improved varieties with higher productivity, nutrients, grain mold tolerance, and climate resilience (Dwivedi et al., 2016; Godwin et al., 2019).

An attempt was made in this study to utilize the sources of variation for grain mold tolerance and a protein content present in sorghum landraces, which were evolved under vulnerable conditions with low inputs after the continuous selection by the farmers. To our knowledge, this is the first study to identify stable, high-yielding, drought- and grain mold-tolerant, and nutritionally rich sorghum varieties that were developed from the landraces in the Indian subcontinent. The objectives of the study were to (1) collect the landraces from the southern and central parts of India and identify agronomically superior landraces with a resistance to grain mold disease; (2) develop potential sorghum varieties from the superior landraces; (3) identify high-yielding, grain mold-tolerant genotypes with a stable performance using AMMI and GGE analyses; and (4) determine the nutrient composition [starch, sugar, protein, iron (Fe), and zinc (Zn)] of the developed sorghum varieties.

MATERIALS AND METHODS

Collection and Maintenance of Landraces

A total of 108 landraces were collected from various locations in the southern and central parts of India in 2008. These landraces represented a diversity for grain maturity, grain color, panicle shape, grain yield, fodder yield, porridge making quality, fodder quality, and tolerance to grain mold disease under field conditions (Supplementary Table 1). In the field evaluation of these landraces from 2008 to 2010, a single-plant selection was followed by self-pollinating main panicles of individual landrace collections for three generations to bring a genetic uniformity within the landraces at the Regional Agricultural Research Station (RARS), Palem, Telangana (Former Andhra Pradesh), India. Based on the construction of passport data and dendrogram (Supplementary Figures 1, 2) using the distances matrix obtained by an unweighted pair group method with arithmetic mean (UPGMA), 36 diverse landraces (PSLRC 1–PSLRC 36) distinct for various characters, viz., maturity, grain type, tolerances to grain mold disease, and terminal moisture stress, were maintained for future breeding (Table 1).

Development of Sorghum Varieties

From 2010 to 2015, hybridization followed by a selection was carried out in rainy and post-rainy seasons by utilizing 11 agronomically superior and grain mold-resistant landraces. A minimum population of 250 plants was maintained in each F₂ and subsequent generation. They were advanced to F₆ generation

by using the pedigree method of selection. The F₆ progeny of individual cross combination was considered as a single-sorghum advanced genotype having a diverse genetic background for agronomic and grain characters and grain mold resistance (Figure 1). About 19 superior advanced sorghum cultivars were identified after the evaluation for two consecutive seasons (rainy and post-rainy seasons, 2014) in an advanced variety trial at RARS, Palem, Telangana, India, among which seven genotypes were characterized with a yellow pericarp and six genotypes were with a brown pericarp. The remaining six genotypes had a white grain and a white grain with black glume (Table 2, Supplementary Figures 3, 4).

Nutritional Composition Analysis

The nutritional composition analysis of 19 sorghum varieties evaluated at E1 (Palem) in 2018 was performed at MFPI-Quality Control Laboratory, Prof. Jayashankar Telangana State Agricultural University, Hyderabad, Telangana, India. Whole grains were collected from the fields where they were grown and analyzed for the protein, total starch, sugar, Fe, and Zn content. The protein level was quantified by using the generic combustion method of analysis with the LECO F-528 nitrogen analyzer (LECO, St. Joseph, MI, USA) calibrated with ethylenediaminetetraacetic acid (EDTA) according to the association of official analytical chemists method (AOAC, 2016). The grain samples were ground to a suitable fineness to pass No. 20 sieve and dried at 102 ± 2°C for 2 h. A moisture-free sample weighing 200 mg was analyzed to estimate protein content. Analyses for total starch, sugar, Fe, and Zn were performed according to Shegro et al. (2012). Starch content was determined by using a total starch assay procedure. The total sugar content in stalks at physiological maturity was estimated as the total soluble sugars by using a handheld refractometer. For the determination of Fe and Zn contents, sorghum grains were ground to a fine powder. About 2-g flour samples were oven-dried for 3 h after which the samples were triple acid digested by the addition of 1 ml of 55% (v/v) HNO₃ after cooling. The acid was evaporated by using a sand bath and the samples were oven-dried again. The samples were moistened by using 10 ml of 55% HNO₃ (1:2 v/v), and they were then placed in the sand bath for 5–10 min. The samples were allowed to dissolve overnight to extract the minerals accordingly.

Evaluation of Varieties for Yield Performance and Screening for Grain Mold Tolerance

About 19 advanced sorghum varieties along with a popular variety (CSV 31) were evaluated over the three rainy seasons (June–September) from 2016 to 2018 at 13 locations for yield and yield-related characters (Table 3). A single location tested for three consecutive seasons of 2016, 2017, and 2018 was considered as one environment. All 13 environments are drought-prone environments characterized by intermittent dry spells with E4, E6, E8, E9, and E11 receiving an annual rainfall of ≤600 mm (Table 3). Each sorghum genotype was planted on six rows of 5-m length plot by using between- and within-row spacing

TABLE 1 | Description of the superior sorghum landraces used in this study.

S. No.	Local accession Number	Sorghum race	Days to 50% flowering	Plant Height (cm)	Glume cover-age %	Grain color	Panicle compactness	Grain mold resistance
1	PSLRC 1	Durra	65	285	71	Yellow	Semi compact	Tolerant
2	PSLRC 2	Durra	68	315	83	Yellow	Semi compact	Tolerant
3	PSLRC 3	Guinea	68	246	79	White	Semi compact	Tolerant
4	PSLRC 4	Durra	66	220	67	Yellow	Semi compact	Tolerant
5	PSLRC 5	Guinea	71	264	54	White	Compact	Susceptible
6	PSLRC 6	Durra	69	310	63	Yellow	Semi compact	Tolerant
7	PSLRC 7	Durra	68	240	71	Yellow	Semi compact	Tolerant
8	PSLRC 8	Durra	65	298	90	Brown	Semi compact	Tolerant
9	PSLRC 9	Durra	71	210	85	Brown	Semi compact	Tolerant
10	PSLRC 10	Durra	72	218	46	Brown	Semi compact	Tolerant
11	PSLRC 11	Durra	69	195	62	Brown	Semi compact	Susceptible
12	PSLRC 12	Durra	68	226	76	Yellow	Semi compact	Tolerant
13	PSLRC 13	Durra	65	315	88	Yellow	Semi compact	Tolerant
14	PSLRC 14	Durra	62	242	77	Yellow	Semi compact	Tolerant
15	PSLRC 15	Durra	64	292	90	Brown	Semi compact	Tolerant
16	PSLRC 16	Durra	66	262	72	Yellow	Semi compact	Tolerant
17	PSLRC 17	Durra	65	245	75	Brown	Semi compact	Tolerant
18	PSLRC 18	Durra	64	272	80	Yellow	Semi compact	Tolerant
19	PSLRC 19	Durra	71	282	69	Yellow	Compact	Tolerant
20	PSLRC 20	Guinea	78	265	54	Black glume	Loose	Tolerant
21	PSLRC 21	Guinea	76	272	78	Black glume	Loose	Tolerant
22	PSLRC 22	Guinea	64	218	43	Black glume	Loose	Tolerant
23	PSLRC 23	Durra	75	234	68	Brown	Semi compact	Tolerant
24	PSLRC 24	Durra	72	240	75	Brown	Semi compact	Tolerant
25	PSLRC 25	Durra	76	265	56	Brown	Loose	Tolerant
26	PSLRC 26	Durra	75	245	70	Yellow	Semi compact	Tolerant
27	PSLRC 27	Durra	68	262	38	Brown	Compact	Tolerant
28	PSLRC 28	Durra	62	275	81	Yellow	Semi compact	Tolerant
29	PSLRC 29	Guinea	65	260	75	White	Compact	Tolerant
30	PSLRC 30	Guinea	66	235	68	White	Semi compact	Tolerant
31	PSLRC 31	Durra	72	292	91	Brown	Loose	Tolerant
32	PSLRC 32	Durra	70	210	94	Yellow	Semi compact	Tolerant
33	PSLRC 33	Durra	68	228	79	Yellow	Semi compact	Tolerant
34	PSLRC 34	Durra	65	235	82	Yellow	Semi compact	Tolerant
35	PSLRC 35	Durra	65	242	69	Yellow	Semi compact	Tolerant
36	PSLRC 36	Durra	65	262	78	Yellow	Semi compact	Tolerant

The above landraces were selected to develop varieties after diversity analysis of 108 landraces following an unweighted pair group method with arithmetic mean (UPGMA) (Supplementary Table 1 and Figure 1).

of 45 and 10 cm, respectively. Weeds, insect pests, and foliar disease control were carried out as recommended for the crop by using a combination of chemical and cultural practices. During harvest, the four central rows within each plot were sampled for grain yield and fodder yield. The 19 genotypes along with susceptible (SPV 462) and resistant (IS 8545) checks were also evaluated in the sorghum grain mold nursery over the three rainy seasons (June–September) in 2016 to 2018 at three locations *viz.*, RARS, Palem, Agricultural Research Station, Tandur, Agricultural Research Station, Madhira, Telangana, India under natural epiphytotic conditions for a grain mold evaluation.

Each genotype was sown in six rows of 5 m in length during the first fortnight of June so that the grain maturity stage coincided with the periods of frequent rainfall received in the ensuing August–September, thus predisposing the crop to grain mold disease. During rain-free days, high relative humidity (>90%) was maintained from the flowering to the physiological maturity stage by using sprinkler irrigation. About 10 uniformly flowering plants with the same flowering window were tagged in each row. The visual panicle grain mold rating (PGMR) was taken on each of the tagged plants at the prescribed physiological maturity by using a progressive 1–9 scale, where 1 = no mold infection,

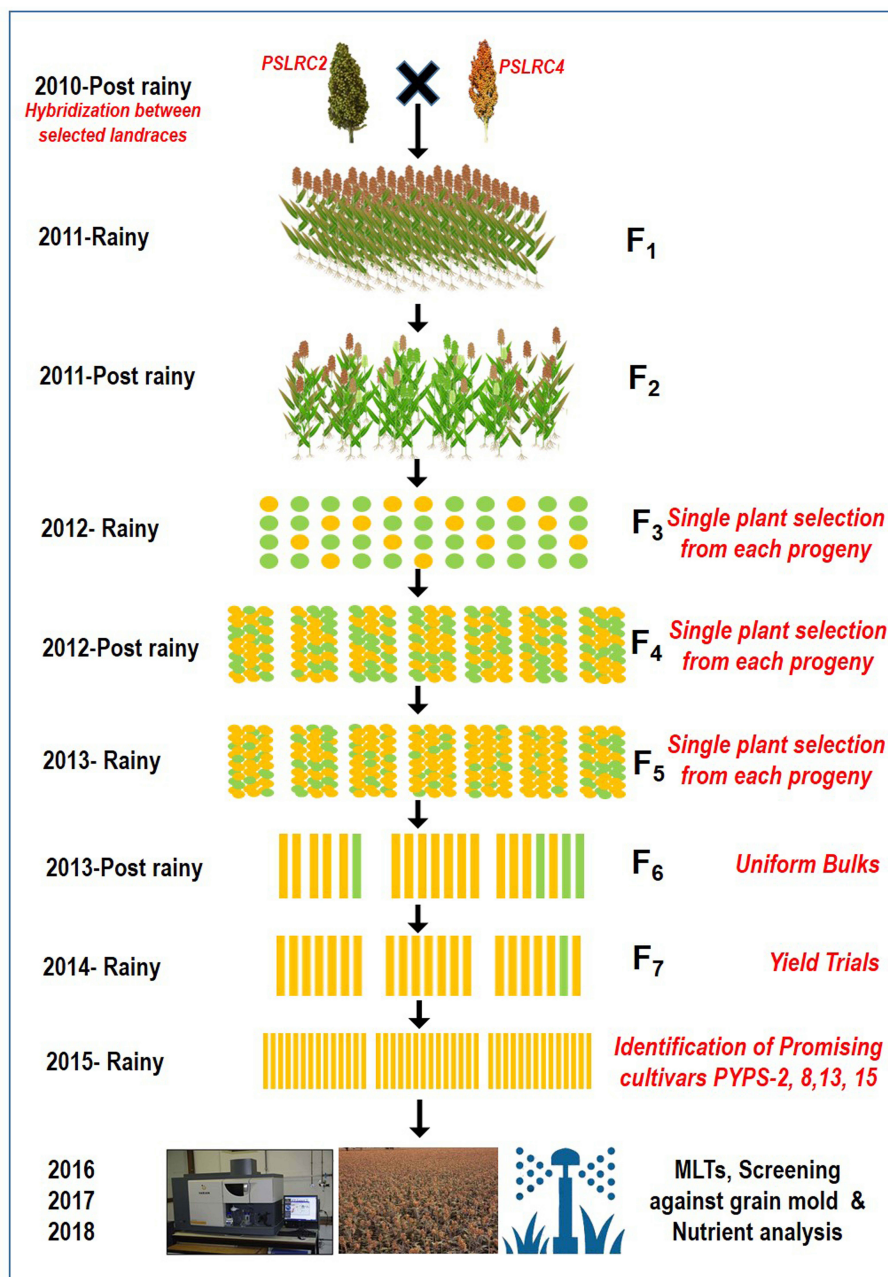


FIGURE 1 | A flow diagram showing the development of sorghum varieties utilizing landraces through a pedigree selection (2010–2015) followed by multi-environment trials (2016–2018).

2 = 1–5%, 3 = 6–10%, 4 = 11–20%, 5 = 21–30%, 6 = 31–40%, 7 = 41–50%, 8 = 51–75%, and 9 = 76–100% molded grains on a panicle (Singh and Bandyopadhyay, 2000; Thakur et al., 2007). All the trials at each location were conducted in a complete randomized block design with three replications.

Data Analysis

Combined ANOVA was performed for yield and disease reaction at 13 and 3, environments, respectively. Statistical computations

and estimations were carried out by using GenStat 18.0 (Goedhart and Thissen, 2010) to partition the yield variation into environments, GGE. The grain yield, fodder yield, and disease resistance reaction data were subjected to the AMMI and GGE biplot analysis. The AMMI model combines both additive and multiplicative components of two-way data structures, which helps in the prediction of potential genotypes and an environmental effect on them (Gauch and Zobel, 1996; Gauch, 2013). The GGE biplots were constituted from the first two

principal components (PC1 and PC2) derived by subjecting the environment-centered yield data (which contains GGE) to singular-value decomposition (SVD). The model for a GGE biplot (Yan et al., 2000) based on SVD of the first two principal components (PC) is:

$$Y_{ij} = \mu + \beta_j + \sum_{k=1}^k \lambda_k \gamma_{ik} \delta_{jk}$$

where Y_{ij} is the mean of genotype i in environment j ; μ is the grand mean; β_j is the environment main effect; k is the number of PC required for appropriate depiction of GGE, n is the singular

value; λ_k is the singular value of the k th PC (PC_k). γ_{ik} and δ_{jk} are the scores of i th genotype and j th environment, respectively, for PC_k .

The GGE biplot software was used to generate graphs showing (1) a “which-won-where” pattern to identify mega-environments, (2) ranking of varieties based on yield and stability, and (3) a correlation of vectors between the environments as per the method described by Yan and Kang (2003).

RESULTS

AMMI Analysis

The combined ANOVA analysis showed highly significant ($p \leq 0.05$) genotype differences over locations and seasons suggesting that both grain and fodder yields varied across the environments. Highly significant environments, genotypes, and $G \times E$ interaction explained 35.3, 23.6, and 29.8% of the total sum of squares for grain yield and 28.9, 23.4, and 25.0% for fodder yield, respectively (Table 4). The magnitude of the environments (E) and $G \times E$ interaction sum of squares were twice larger than that for genotypes sum of squares indicating ample of variations in the genotypic response across the environments for both grain yield (58.9%) and fodder yield (53.9%). Further partitioning of the $G \times E$ interaction sum of squares resulted in two significantly interaction PC axes (IPCA1 and IPCA2), which explained 35.9 and 20.7% of the variation, respectively, and together contributed to 56.6% of the total $G \times E$ interaction for grain yield. Similarly, for fodder yield, IPCA1 and IPCA2 explained 54.9 and 11.9% of the $G \times E$ interaction, respectively, and together contributed to 66.8% of the total variation. This explained the differential performance of genotypes for grain yield and fodder yield across the environments.

AMMI Stability Value

The AMMI stability value (ASV) proposed by Purchase et al. (2000) is used to identify stable genotypes and environments. For environments and genotypes, a low ASV indicates that the environments and genotypes are highly stable and least

TABLE 2 | Sorghum genotypes evolved through hybridization followed by a selection from the superior landraces.

S. No.	Genotype	Cross combination	Grain color
1.	PYPS-1	PSLRC 2 × PSLRC 3	Yellow
2.	PYPS-2	PSLRC 2 × PSLRC 4	Yellow
3.	PYPS-3	PSLRC 3 × PSLRC 4	White
4.	PYPS-4	PSLRC 2 × PSLRC 6	Yellow
5.	PYPS-5	PSLRC 21 × PSLRC 7	White grain with black glume
6.	PYPS-6	PSLRC 3 × PSLRC 6	White
7.	PYPS-7	PSLRC 3 × PSLRC 7	White
8.	PYPS-8	PSLRC 8 × PSLRC 9	Brown
9.	PYPS-9	PSLRC 8 × PSLRC 10	Brown
10.	PYPS-10	PSLRC 9 × PSLRC 10	Brown
11.	PYPS-11	PSLRC 4 × PSLRC 6	Yellow
12.	PYPS-12	PSLRC 4 × PSLRC 12	Yellow
13.	PYPS-13	PSLRC 2 × PSLRC 21	Yellow
14.	PYPS-14	PSLRC 20 × PSLRC 21	White grain with black glume
15.	PYPS-15	PSLRC 2 × PSLRC 7	Yellow
16.	PYPS-16	PSLRC 20 × PSLRC 7	White grain with black glume
17.	PYPS-17	PSLRC 2 × PSLRC 8	Brown
18.	PYPS-18	PSLRC 3 × PSLRC 8	Brown
19.	PYPS-19	PSLRC 4 × PSLRC 8	Brown

TABLE 3 | Details of the 13 environments tested for yield and stability of 19 sorghum genotypes in Telangana, India.

Environment	Location	Latitude	Longitude	Soil type	Rainfall (mm)
E1	Palem	16.5461° N	78.2077° E	Red sandy	690
E2	Tandur	17.2576° N	77.5875° E	Sandy loam	780
E3	Madhira	16.9182° N	80.3633° E	Sandy loam	750
E4	Hanwada, Mahabubnagar	16.8106° N	77.9196° E	Red sandy	600
E5	Kodangal	17.1103° N	77.6235° E	Sandy loam	760
E6	Gaddamallaihuda	17.0974° N	78.6867° E	Red sandy	560
E7	PA Pally	16.6996° N	79.0267° E	Sandy loam	700
E8	Maddur	15.8563° N	77.2431° E	Sandy loam	600
E9	Aler	17.6437° N	79.0430° E	Red sandy	580
E10	Kulkacherla	17.0161° N	77.8746° E	Red sandy	630
E11	Ramapuram	15.9653° N	77.9410° E	Red sandy	580
E12	Kothakota	16.3787° N	77.9410° E	Red sandy	720
E13	Devarakadara	16.6248° N	77.8410° E	Red sandy	650

TABLE 4 | Additive main effects and multiplicative interaction (AMMI) ANOVA for grain yield and fodder yield of 19 sorghum genotypes evolved from landraces over 13 locations in 3 years (2016–2018).

Source	DF	Grain Yield					Fodder Yield			
		SS	MS	F	% contribution SS		SS	MS	F	% contribution SS
Total	2,222	1,087,099,555	489,244	–			36,509,998,551	16431142.46		
Genotypes	18	324,366,092	18,020,338**	520.53	23.6		8,547,745,363	474874742.38**	200.98	23.4
Environments	12	384,294,081	32,024,506**	925.05	35.3		10,581,124,368	881,760,364**	373.20	28.9
Blocks	26	53,982,701	2,076,257	59.97	4.96		2,582,926,484	99343326.30	420.85	7.07
Interaction	216	256,949,082	1,189,580**	34.4	29.8		9,127,499,637	42256942.76**	17.88	25.0
IPCA1	29	92244720.43	3180852.42	91.88	35.9		5010997300.71	172793010.36	73.13	54.9
IPCA2	27	53,188,460	1,969,943	56.90	20.7		1086172456.80	40228609.51	170.24	11.9
Residuals	160	111515901.6	96,356	3221.23	43.7		3,030,329,880	18,939,562	1282.58	33.2
Error	1,950	67,507,598	34,619	–			4,607,225,870	2362679.93		

**significant at 1% probability level.

DF, Degrees of freedom; SS, Sum of squares; MS, Mean sum of squares; F, F-calculated value; IPCA, Interaction principal component axis.

interactive whereas a high ASV indicates that the environments and genotypes are highly interactive and unstable. Based on the ASV for grain yield, E5, E11, and E13 were the most stable and high yielding environments (Table 5). On the contrary, E4 followed by E10 was the most unstable and most interactive environment with high ASV scores for grain yield. For fodder yield, the environments E6 and E10 were mostly stable with low ASV scores, and the environments E5, E1, and E4 were mostly unstable with high ASV scores (Table 5).

Sorghum varieties PYPS 2, PYPS 5, PYPS 8, PYPS 13, PYPS 14, and PYPS 17 were the most stable varieties for grain yield, and PYPS 11, PYPS 16, PYPS 7, PYPS 4, and PYPS 18 were the most unstable varieties (Table 6). For fodder yield, the varieties PYPS 16, PYPS 10, PYPS 13, PYPS 15, and PYPS 2 were the most stable, and PYPS 8, PYPS 5, and PYPS 9 were unstable (Table 6).

Genotype Stability Index

Genotype stability index (GSI) can be used to classify stable genotypes incorporating both yield and stability in a single non-parametric index (Singh et al., 2019). The GSI considered the ranks of the genotype yield across the environments and ASVs. GSI was calculated as the sum of Rank of ASV [RASV (ASV)] and RY (Rank of mean genotype yield of all environments). Considering high grain and fodder yields, a high protein content and moderate resistance to grain mold, low ASV values and high GSI, sorghum varieties PYPS 1 and PYPS 13 were identified as the best stable genotypes. Sorghum varieties PYPS 2, PYPS 8, PYPS 12, PYPS 15, and PYPS 16 having high grain and fodder yields with either a low ASV or a high GSI were also known as the best stable varieties across the environments (Table 6).

AMMI Biplot Analysis

In the AMMI1 and AMMI2 biplots (Figure 2), the environments were designated by the letter “E” followed by numbers 1–13 as suffix while the genotypes were represented by the letter “G” followed by numbers 1–19. In the AMMI1 biplot, the main effects (genotype mean and environment mean) on abscissa were

plotted against the respective IPCA1 scores on the ordinate (Yan et al., 2007). The quadrants (Q) in the AMMI1 graph represent: higher mean (QI and QII), lower mean (QIII and QIV), +ve IPCA1 score (QI and QIV), and –ve IPCA1 score (QII and QIII). When a genotype and an environment have the same sign on the IPCA1 axis, their interaction is positive and, if opposite, their interaction is negative. Thus, if a variety has a IPCA1 score nearer to zero, it has a small interaction effect and was considered as stable over wide environments. On the other hand, genotypes with high mean yield and large IPCA1 scores were considered as explicitly adapted to specific environments (Abdi and Williams, 2010).

Accordingly, in the present study, sorghum varieties PYPS 2, PYPS 8, and PYPS 15 were specifically adapted to the high-yielding environments E1 and E10 and the varieties PYPS 13 and PYPS 16 were adapted to the environments E5 and E9 with grain yield more than the grand average yield (Figure 2A). Furthermore, the varieties PYPS 2, PYPS 8, and PYPS 13 were also more stable in comparison to PYPS 4, PYPS 7, PYPS 11, and PYPS 16. Similarly, the varieties PYPS 1, PYPS 5, PYPS 6, PYPS 10, and PYPS 14 were nearer to zero indicating that they are highly stable for grain yield than other varieties. For fodder yield, the varieties PYPS 2, PYPS 13, and PYPS 8 were specifically adapted to high-yielding environments E1, E5, E9, and E10 (Figure 2B). The varieties PYPS 10, PYPS 12, PYPS 15, and PYPS 16 were more stable in comparison to PYPS 4, PYPS 5, PYPS 9, PYPS 17, and PYPS 8 as these genotypes were far from the origin. The varieties PYPS 11, PYPS 15, and PYPS 16 were nearer to zero indicating a higher stability for fodder yield than the other genotypes.

The AMMI2 biplot is a graphical representation of the interaction effect wherein the relationship between the genotypes and environments is depicted in a vector view (Guerra et al., 2009). The biplot detects the environments and genotypes that contributed least to the interaction (most stable) as well as the desirable combinations of genotypes and environments in terms of specific adaptability. The statistically stable genotypes and

TABLE 5 | Mean performance for grain yield, fodder yield, IPCA1, IPCA2 scores, and ASV values of 13 environments.

S. No.	Environment	Grain yield (kg/ha)	IPCA1	IPCA2	ASV	Fodder yield (kg/ha)	IPCA1	IPCA2	ASV
1	E1	3,398	−10.36	−1.58	17.96	15,970	−41.17	1.82	189.00
2	E2	2,860	−8.68	−6.32	16.27	13,893	−20.58	−2.80	94.30
3	E3	2,707	2.90	16.71	17.40	12,707	34.07	5.56	156.80
4	E4	2,417	18.31	−10.83	33.40	11,705	37.63	−23.84	174.60
5	E5	3,396	1.85	2.54	4.06	15,968	−43.31	−5.13	199.06
6	E6	2,883	−11.63	−5.58	16.00	13,878	0.88	−15.41	15.90
7	E7	2,716	0.57	13.30	13.30	12,776	14.04	17.09	66.70
8	E8	2,699	10.02	0.01	17.32	12,931	16.29	−8.10	75.20
9	E9	3,076	6.07	−11.80	15.80	14,591	−8.47	−31.76	50.20
10	E10	3,044	−18.98	−4.60	32.30	14,314	−5.27	7.79	25.40
11	E11	2,865	−2.97	7.46	9.00	13,844	−17.66	17.40	82.84
12	E12	2,659	6.72	6.79	13.40	12,852	21.44	3.69	98.50
13	E13	2,725	6.20	−6.10	12.30	12,996	12.10	33.71	57.00
	Overall mean	2880.38				13725.10			
	LSD (0.05)	341.45				1159.50			

IPCA, Interaction principal component axis; ASV, AMMI stability value; LSD, Least significant difference.

environments are represented by the points nearer to the origin in the AMMI2 biplot, with the values being nearer to zero for the two axes of interaction (IPCA1 and IPCA2). The discrimination power of a test environment is proportional to the length of the environment vector, which is the line connecting the origin and test environment and those genotypes falling apart from the origin with long spokes were termed as highly interacting genotypes (Yan and Holland, 2010).

In this study, for grain yield, the environments E4, E3, E10, and E9 were farthest from the origin and were the most discriminating but non-representative (unstable) while E5, E11, E2, E6, and E13 lied closest to the origin and contributed least to the $G \times E$ interaction (**Figure 3A**). They were the most representative (stable) environments but with poor discriminating ability. Sorghum varieties PYPS 11, PYPS 15, PYPS 10, and PYPS 7 were more responsive since they were away from the origin, whereas PYPS 5, PYPS 2, PYPS 17, and PYPS 14 were closer to the origin, and hence they were less sensitive to environmental changes for grain yield (**Figure 3A**). For fodder yield, the environments E13, E9, E4, E5, and E1 were the most discriminating but non-representative (unstable) and E10, E8, and E6 were the most stable environments (**Figure 3B**). Sorghum varieties PYPS 13, PYPS 11, PYPS 8, and PYPS 2 were more responsive and the varieties PYPS 19, PYPS 16, PYPS 1, PYPS 14, and PYPS 3 were less sensitive to changes in the environment (**Figure 3B**).

GGE Biplot Analysis

The GGE biplot developed by Yan et al. (2000) displays the genotype main effect (G) plus $G \times E$ interaction, which are the two sources of variation that are relevant to a cultivar evaluation. The which-won-where pattern first described by Yan et al. (2000) identifies the best performer for a site(s) and defines mega-environments (subregions) by selecting the superior genotypes

for each mega-environment, thus effectively exploiting both genotypes and $G \times E$ interaction.

(a) The which-won-where pattern

The polygon view of the GGE biplot displays the “which-won-where” pattern by connecting the markers of the genotypes that are further away from the biplot origin such that all the other genotypes are contained in the polygon (Yan et al., 2000). Genotypes having the specific adaptive ability for a specific environment or group of environments were identified by using this pattern. The biplot is further divided into sectors delimited by the lines perpendicular to each side of the polygon. The genotypes in a sector are similar in performance compared to the genotypes in other sectors.

In the present study, the biplot is divided into five sectors for grain yield (**Figure 4A**) and four sectors for fodder yield (**Figure 4B**). The varieties PYPS 2, PYPS 4, PYPS 5, PYPS 7, PYPS 11, and PYPS 16 were situated at the apex of the polygon, representing the highest grain yield and indicated superior genotypes (**Figure 4A**). Sorghum varieties PYPS 1, PYPS 9, PYPS 14, and PYPS 17 were closest to the center of origin indicated a low variation in the $G \times E$ interaction for grain yield (**Figure 4A**). The variety PYPS 16 was suitable for the three environments E4, E8, and E12. The varieties PYPS 2, PYPS 8, PYPS 13, and PYPS 15 were suitable for the remaining 10 environments E1, E2, E3, E5, E6, E7, E9, E10, E11, and E13 for grain yield (**Figure 4A**). The varieties PYPS 2, PYPS 5, PYPS 6, PYPS 10, PYPS 13, and PYPS 14 were situated at the apex of the polygon, representing the highest fodder yield and indicated superior genotypes (**Figure 4B**). Sorghum variety PYPS 2, PYPS 15, and PYPS 1 were suitable for E2, E3, E4, E6, E7, E8, E10, E12, and E13 for fodder yield. The varieties PYPS 8, PYPS 13, and PYPS 11 were suitable for environments E1, E2, E5, E9, and E11 (**Figure 4B**). The varieties PYPS 1, PYPS 3, PYPS 7, and PYPS 11 were closest to the center of the origin indicated a low variation in the $G \times E$ interaction for fodder yield (**Figure 4B**).

TABLE 6 | Classification of stable sorghum varieties based on mean performance, ASV and stability index for grain yield, fodder yield, disease reaction, and protein.

S. No.	Genotype	Grain yield (kg/ha) and its rank	IPCA1	IPCA2	ASV and RASV	GSI and RGSi	Fodder yield (kg/ha) and its rank	IPCA1	IPCA2	ASV and RASV	GSI and RGSi	Disease Score and rank	ASV and RASV	GSI and RGSi	Protein (%) and rank
1	PYPS-1	2,756 (7)	−0.36	5.56	5.6 (7)	13 (4)	14,108 (6)	−7.79	−1.13	35.8 (9)	14 (5)	4.52 (10)	0.01 (1)	11 (2)	12.32 (4)
2	PYPS-2	3,698 (1)	−2.46	0.96	4.4 (3)	4 (1)	20,586 (1)	−23.99	20.70	18.6 (7)	8 (2)	3.92 (3)	0.6 (18)	21 (14)	14.52 (1)
3	PYPS-3	2,639 (14)	−6.14	−5.09	11.7 (12)	25 (13)	12,448 (12)	7.95	1.06	36.5 (10)	21 (10)	4.5 (8)	0.11 (10)	18 (10)	9.78 (15)
4	PYPS-4	2,603 (18)	−12.54	2.18	21.8 (16)	33 (18)	13,008 (9)	20.84	6.84	96 (16)	24 (13)	4.1 (5)	0.4 (16)	21 (15)	10.3 (14)
5	PYPS-5	2,514 (20)	−1.09	−1.00	2.1 (1)	20 (9)	11,793 (16)	26.90	16.18	124.7 (18)	33 (19)	4.51 (9)	0.09 (7)	16 (7)	10.52 (12)
6	PYPS-6	2,604 (17)	−2.26	8.07	8.9 (8)	24 (12)	12,335 (13)	−9.22	−14.77	44.8 (12)	24 (13)	4.73 (13)	0.32 (13)	26 (16)	9.63 (16)
7	PYPS-7	2,575 (19)	11.81	13.00	24.2 (17)	35 (19)	13,059 (8)	7.76	−12.63	37.8 (11)	18 (7)	4.18 (6)	0.18 (11)	17 (8)	9.25 (18)
8	PYPS-8	3,584 (2)	−0.74	−2.43	2.7 (2)	4 (1)	18,632 (2)	−49.35	15.59	227.5 (19)	21 (10)	3.71 (2)	0.04 (3)	5 (1)	13.26 (3)
9	PYPS-9	2,712 (12)	2.64	8.65	9.8 (9)	20 (9)	11,984 (15)	22.22	−12.69	102.8 (17)	31 (17)	4.78 (15)	0.09 (7)	22 (15)	14.13 (2)
10	PYPS-10	2,619 (16)	−0.32	9.98	10.0 (10)	25 (13)	11,538 (18)	2.22	2.37	10.81 (2)	20 (8)	5.1 (18)	0.37 (15)	33 (19)	9.58 (17)
11	PYPS-11	2,699 (13)	−18.36	1.28	31.8 (19)	31 (17)	13,895 (7)	9.28	26.82	50.3 (14)	20 (8)	3.65 (1)	0.23 (12)	13 (5)	10.73 (8)
12	PYPS-12	2,729 (9)	−6.61	−2.17	11.6 (11)	18 (7)	12,728 (10)	0.83	−15.84	16.2 (5)	14 (5)	4.8 (16)	0.06 (4)	20 (12)	10.95 (6)
13	PYPS-13	3,514 (3)	0.88	−4.65	4.6 (4)	7 (3)	18,122 (3)	−2.55	−2.00	11.8 (3)	6 (1)	3.97 (4)	0.09 (7)	11 (4)	10.66 (10)
14	PYPS-14	2,718 (10)	1.90	−3.31	4.7 (5)	14 (6)	11,392 (20)	3.13	8.36	16.6 (6)	25 (15)	5.32 (19)	0.07 (5)	14 (6)	10.52 (12)
15	PYPS-15	3,432 (4)	−3.74	−14.39	14.4 (14)	18 (7)	16,955 (4)	−0.20	−12.80	12.8 (4)	8 (2)	4.54 (11)	0.08 (6)	17 (9)	8.78 (19)
16	PYPS-16	3,268 (5)	17.02	−10.70	31.3 (18)	23 (11)	12,561 (11)	1.01	0.98	4.74 (1)	11 (4)	5 (17)	0.02 (2)	19 (11)	10.77 (7)
17	PYPS-17	2,729 (8)	2.80	1.21	4.9 (6)	13 (4)	11,697 (17)	16.07	14.06	75.3 (15)	31 (17)	4.74 (14)	0.5 (17)	31 (18)	11.28 (5)
18	PYPS-18	2,713 (11)	11.86	−3.92	20.8 (15)	25 (13)	11,627 (19)	9.62	−13.32	46.15 (13)	30 (16)	4.73 (12)	0.72 (19)	31 (17)	10.6 (11)
19	PYPS-19	2,620 (15)	6.98	3.78	12.6 (13)	27 (16)	12,309 (14)	4.23	20.70	27.9 (8)	21 (10)	4.39 (7)	0.32 (13)	20 (13)	10.72 (9)
	CSV 31	2,896 (6)		−3.26			16,588 (5)	−1.62	−11.36						
	Grand mean	2881.1					13868.3								
	LSD (0.05)	412.68					2889.87					2.54			2.15

Value in the parenthesis indicates rank.

IPCA, Interaction principal component axis; ASV, AMMI stability value; RASV, Rank of ASV; GSI, Genotype stability index; RGSi, Rank of GSI; LSD, Least significant difference.

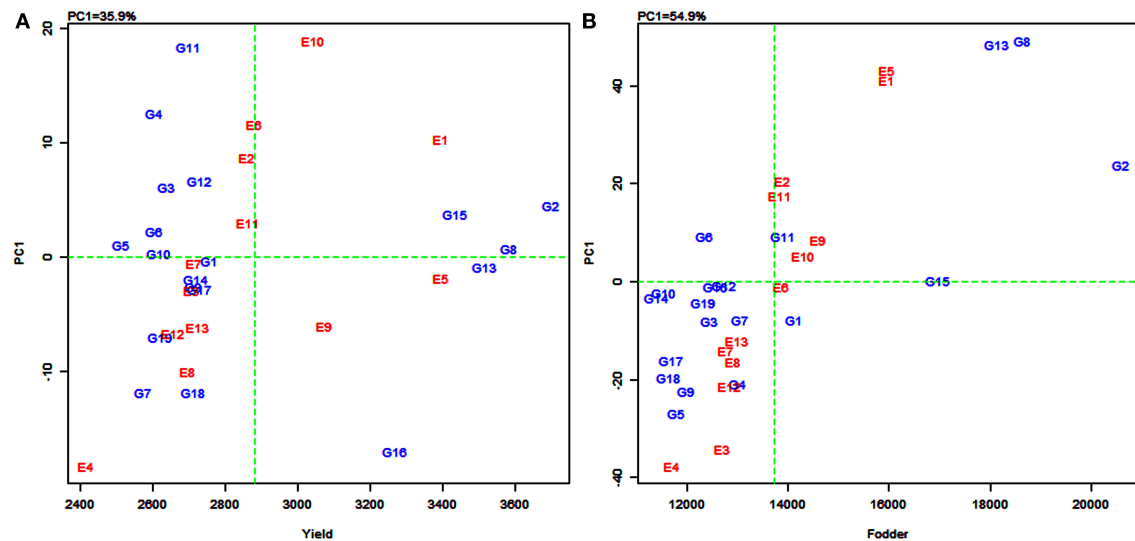


FIGURE 2 | Additive main effects and multiplicative interaction (AMMI1) biplot depicting additive effects vs. interaction principal component axes (IPCA1) for grain yield (A) and fodder yield (B) of 19 sorghum varieties tested across 13 environments.

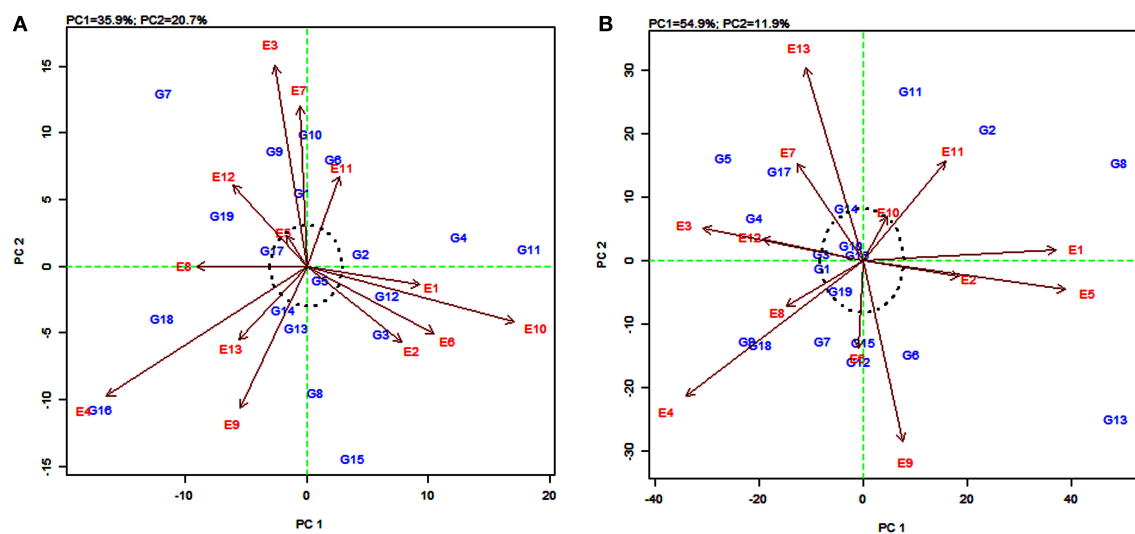


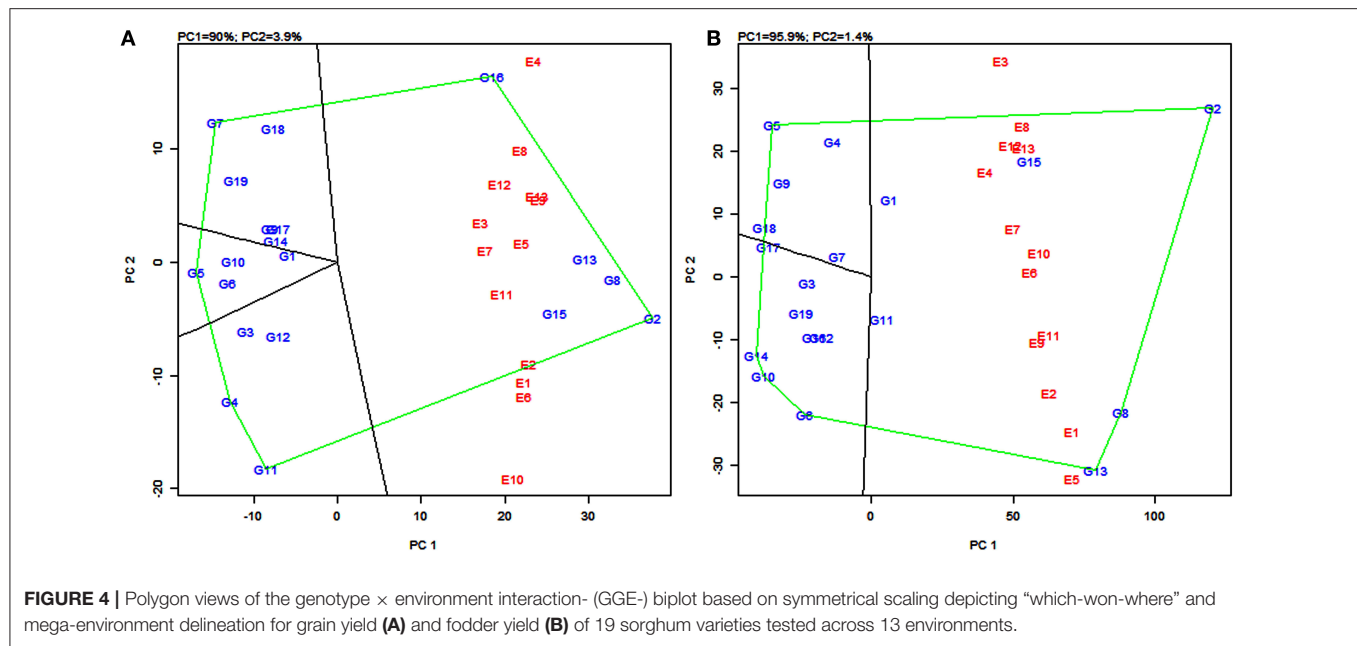
FIGURE 3 | AMMI2 biplot showing two main axes of interaction (IPCA2 vs. IPCA1) for grain yield (A) and fodder yield (B) of 19 sorghum varieties tested across 13 environments.

(b) Mega-environments

In the GGE biplot for grain yield, the five lines (rays) divided the biplot into five sectors. Environments were present in four sectors and these were considered as four mega-environments (Figure 4A) and the superior genotypes for each mega-environment were positioned at the vertex. The GGE biplot for grain yield resulted in four mega-environments *viz.*, first, a mega-environment comprising E8 and E12 with PYPS 16 as the best performing variety; second, a mega-environment comprising E9, E13, E3, E7, E5, E11, and E2 with PYPS 2 performing the best; third, a mega-environment comprising

of E1, E6, and E10, where once again, PYPS 2 was the best performing variety; and fourth, a mega-environment with only one environment E4. Sorghum varieties PYPS 4, PYPS 5, PYPS 7, and PYPS 11 were located at the vertices in the sectors that did not show any environment, indicating that these genotypes were not superior in any of the mega-environments (Figure 4A).

Similarly, in the GGE biplot for fodder yield, four mega-environments were identified *viz.*, the first mega-environment comprising E4, E6, E7, E8, E10, E12, and E13 with PYPS 2 as the best performing variety; the second mega-environment comprising E1, E2, E9, and E11 with PYPS 8 and PYPS 13 as



the best performing varieties (Figure 4B). The third and fourth mega-environment consisted of the single environment of E3 and E5, respectively, suitable for PYPS 2 and PYPS 13. Sorghum varieties PYPS 5, PYPS 6, PYPS 10, and PYPS 14 were located at the vertices in the sectors that did not show any environment, and hence they were not suitable for any mega-environment (Figure 4B).

In addition, the GGE biplot was used to graphically estimate the pattern of environments and discriminate the genotypes (Yan et al., 2000) based on the environment-focused scaling and genotype-focused scaling.

GGE Biplot of Environment View for Yield

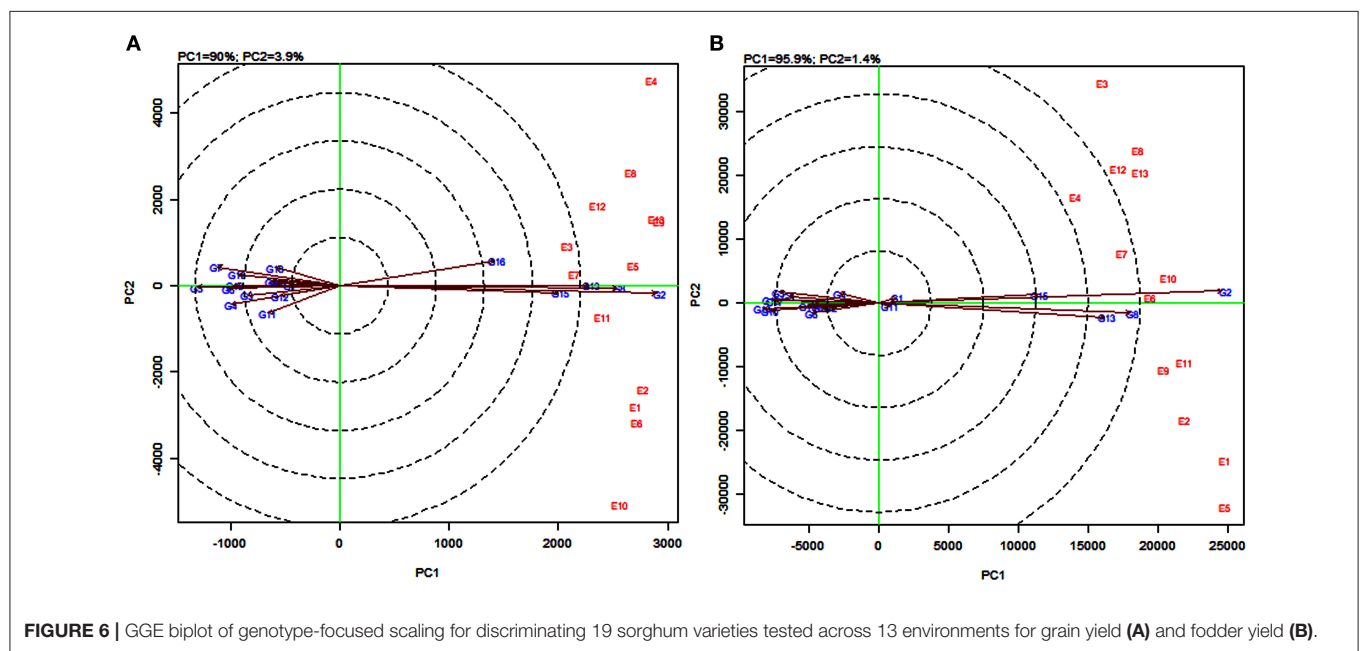
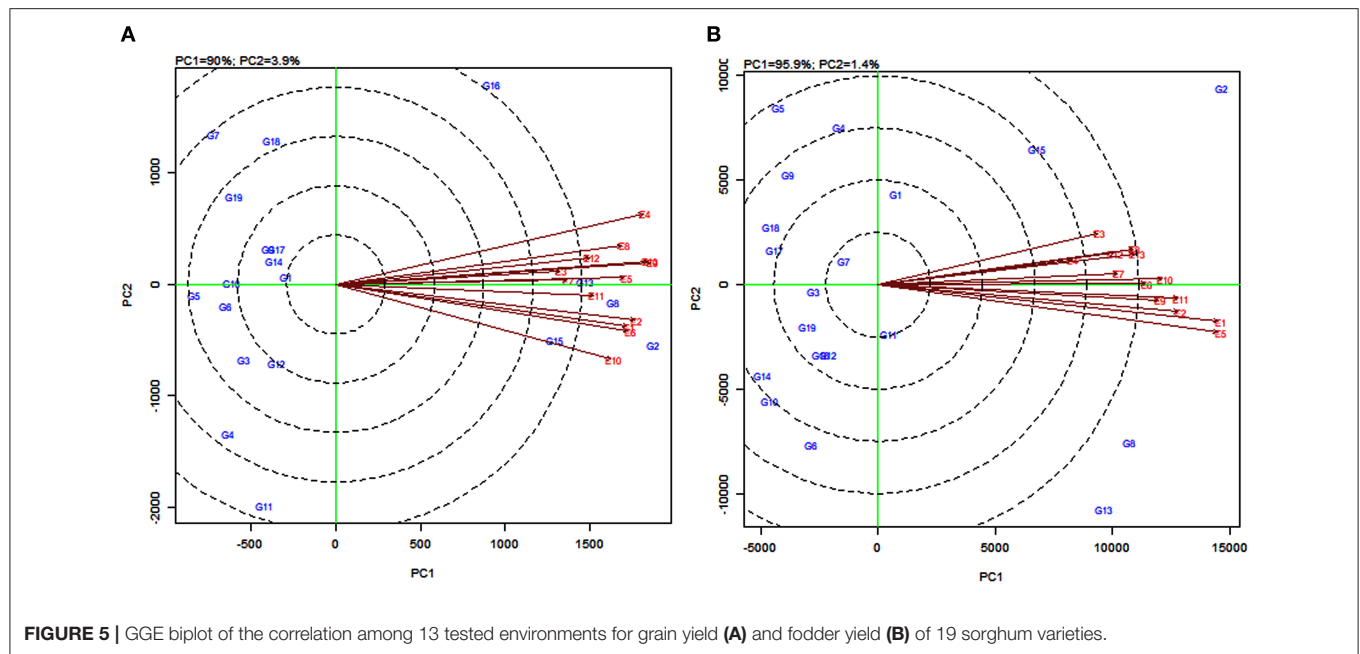
The environmentally centered GGE biplot was used to estimate the pattern of environments for grain yield (Figure 5A) and fodder yield (Figure 5B). To compare the relationship between the environments, lines were drawn to connect the test environments to the biplot origin as environment vectors. The angle cosine between the two environments was used to determine the correlation between them (Dehghani et al., 2010). For grain yield, the angles between the vectors of the majority of the environments were acute, with few overlapping with one another, indicating a positive correlation (Figure 5A). For example, there was an overlapping between the vectors for the environments E3, E9, and E13, and also between E5 and E7. The presence of a wide angle between E4 and E10 indicated that they were negatively correlated and were not similar (Figure 5A). Similarly, for fodder yield, the presence of small angles between the vectors for an environment indicated that the tested environments were similar (Figure 5B). The widest angle between the vector of E3 and E5 suggested a dissimilarity between these two environments.

GGE Biplot of Genotype View for Grain Yield

The vector view of GGE biplot in the genotype-focused scaling measured their dissimilarity in discriminating the genotypes (Kumar et al., 2021). For grain yield, sorghum varieties PYPS1, PYPS 6, PYPS 7, PYPS 9, PYPS 10, PYPS 17, PYPS 18, and PYPS 19 showed the same group position. The varieties PYPS 3, PYPS 4, PYPS 5, PYPS 8, PYPS 11, and PYPS 12 fell in the same group. Likewise, the varieties PYPS 2, PYPS 13, PYPS 14, and PYPS 15 were in the same group. One sorghum variety PYPS 16 with a distinct group was discriminating suggesting dissimilarity with other groups (Figure 6A). For fodder yield, PYPS 1, PYPS 2, and PYPS 15 showed the same group position (Figure 6B). The varieties PYPS 8, PYPS 11, and PYPS 13 showed the same group position with the remaining varieties in two different group positions.

GGE Biplot on Environment for Comparing Environments With an Ideal Environment

Discriminating ability and representativeness of the testing environments are important measures of the GGE biplot. The concentric circles in Figure 7 help us in visualizing the length of the environment vectors, which are a measure of the discriminating ability of the environments and the SD within the respective environments (Yan and Tinker, 2006). The average environment that is drawn as a small circle at the end of the arrow (Figures 7A,B) has the average coordinates of all test environments, and the average environment axis (AEA) is the line passing through the average environment and the biplot origin. A test environment showing a smaller angle with the AEA is more representative than the other test environments (Yan and Rajcan, 2002).



For grain yield, the environments E4, E9, and E13 were the most discriminating genotypes, very closely followed by E1, E2, E6, and E10. The environments E3 and E7, with the shortest vectors from the origin, provided little or no information about the genotype difference and were considered as non-discriminative environments (**Figure 7A**). Based on the angle of the environment with AEA, the environments E5 and E11 were the most representative whereas E4 and E10 were the least representatives for grain yield. For fodder yield, the environments E1 and E5 were the most discriminating genotype while E2, E9, E10, and E11 were moderately

discriminating. The environment E4 followed by E3 and E7 were considered as non-discriminative environments. Further, the environments E10 and E6 were the most representative whereas E3 and E5 were the least representatives for fodder yield (**Figure 7B**). The environments E5, E9, and E13 (**Figure 7A**) and the environments E11 and E12 (**Figure 7B**) located in the first concentric circles were identified as the most ideal environments for obtaining high grain and fodder yields, respectively. The evaluation in these environments maximized the observed genotypic variation among the 19 tested sorghum varieties.

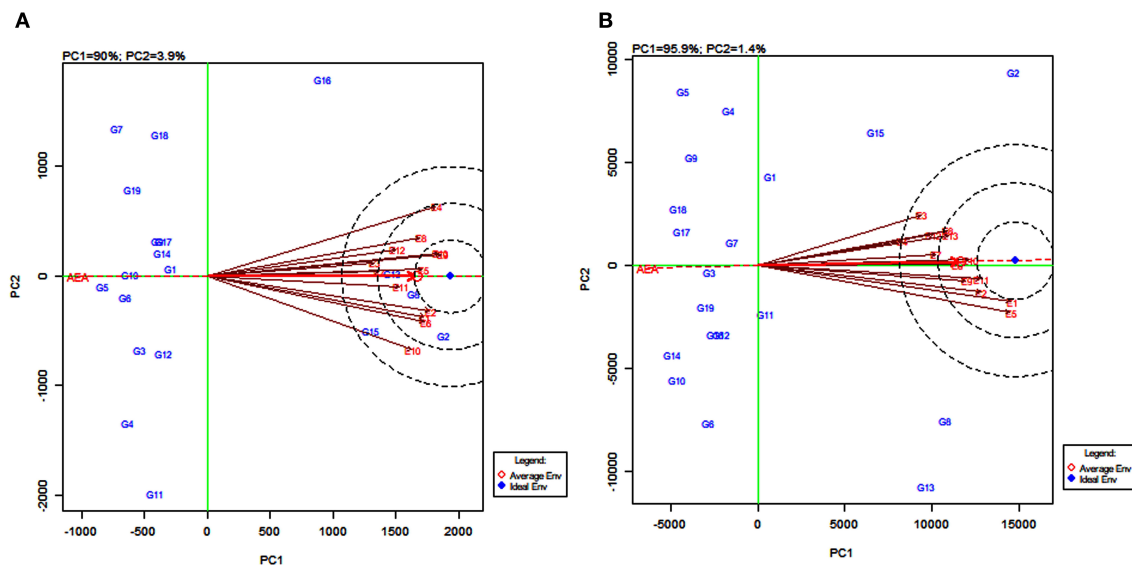


FIGURE 7 | GGE biplot of environment-focused scaling for comparing 13 test environments with an ideal environment for 19 sorghum varieties for grain yield **(A)** and fodder yield **(B)**.

GGE Biplot of Stability and Mean Performance of Genotypes Across Average Environments

The line passing through the biplot origin and the average environment with a single arrow is the AEA. Projections of genotype markers to the AEA depict the mean yield of genotypes (Figure 8). For grain yield, PYPS 2 was the high-yielding variety and PYPS 5 was the lowest yielding variety (Figure 8A). For fodder yield, PYPS 2 was once again the high-yielding variety and PYPS 10 was the lowest yielding variety (Figure 8B). A double arrowed line passing through the biplot origin and perpendicular to the AEA abscissa is the AEA ordinate (Figure 8). A greater projection onto AEA ordinate regardless of the direction means a greater stability. Accordingly, sorghum varieties PYPS 2, PYPS 5, PYPS 8, PYPS 13, PYPS 14, and PYPS 17 with shorter projections over the environments were stable, and the varieties PYPS 11 and PYPS 16 were unstable for grain yield (Figure 8A). For fodder yield, PYPS 2, PYPS 10, PYPS 15, and PYPS 16 were mostly stable over the environments, and the varieties PYPS 5, PYPS 8, and PYPS 13 were unstable (Figure 8B).

AMMI and GGE Analyses for Grain Mold Tolerance

Sorghum varieties PYPS 8 and PYPS 13 recorded low disease scores of 3.71 and 3.97, respectively, for grain mold incidence. Further, these varieties showed low IPCA1 values (0.01 and 0, respectively) and ASVs (0.04 and 0.09, respectively), and hence they were considered as the most stable varieties against grain mold disease (Table 7, Supplementary Table 2). Sorghum variety PYPS 2 showed less incidence of grain mold with a score of 3.92. However, the ASV was relatively high at 0.60, and hence it was moderately stable.

The AMMI biplot revealed that the varieties PYPS 2, PYPS 11, and PYPS 19 had a low grain mold incidence and were moderately stable (Figure 5A). The varieties PYPS 4 and PYPS 7 had low grain mold disease but were unstable. The variety PYPS 14 and the susceptible check SPV 462 were stable with a high grain mold incidence. The GGE biplot graphic analyses of 19 sorghum varieties tested at three environments revealed the magnitude of the interaction of each genotype and environment for grain mold incidence (Figure 5B). PYPS 2, PYPS 4, PYPS 8, and PYPS 11 were the most stable varieties for E2 with a low grain mold incidence followed by the varieties PYPS 7 and PYPS 13. The susceptible check SPV 462, though suitable for environment E2, recorded a high incidence of grain mold.

Nutrient Composition of Sorghum Varieties

The starch, sugar, protein, Fe, and Zn contents varied among the 19 tested sorghum varieties (Table 8). The starch content ranges from 32.11% (PYPS 16) to 57.09% (PYPS 13). The total sugar content among the tested varieties varied from 5.25% (PYPS 9) to 14.93% (PYPS 15). The highest percentage of total protein content was encountered in the grains of PYPS 2 (14.52%) and PYPS 8 (14.13%) whereas the lowest one was demonstrated in PYPS 15 (8.78%). There were significant ($p < 0.05$) differences in the total Fe content among the tested sorghum varieties, with the highest total Fe content in PYPS 4 (12.75 mg) followed by PYPS 2 (10.75 mg) and the lowest in PYPS 15 (4.31 mg) and PYPS 18 (4.40 mg). Significant differences were also found among the sorghum varieties for the total Zn content. The varieties PYPS 1 (3.40 mg), PYPS 8 (3.40 mg), and PYPS 9 (3.30 mg) had the highest Zn content while the varieties PYPS 15 (1.80 mg), PYPS 5 (1.90 mg), PYPS 7 (1.90 mg) recorded the lowest Zn content.

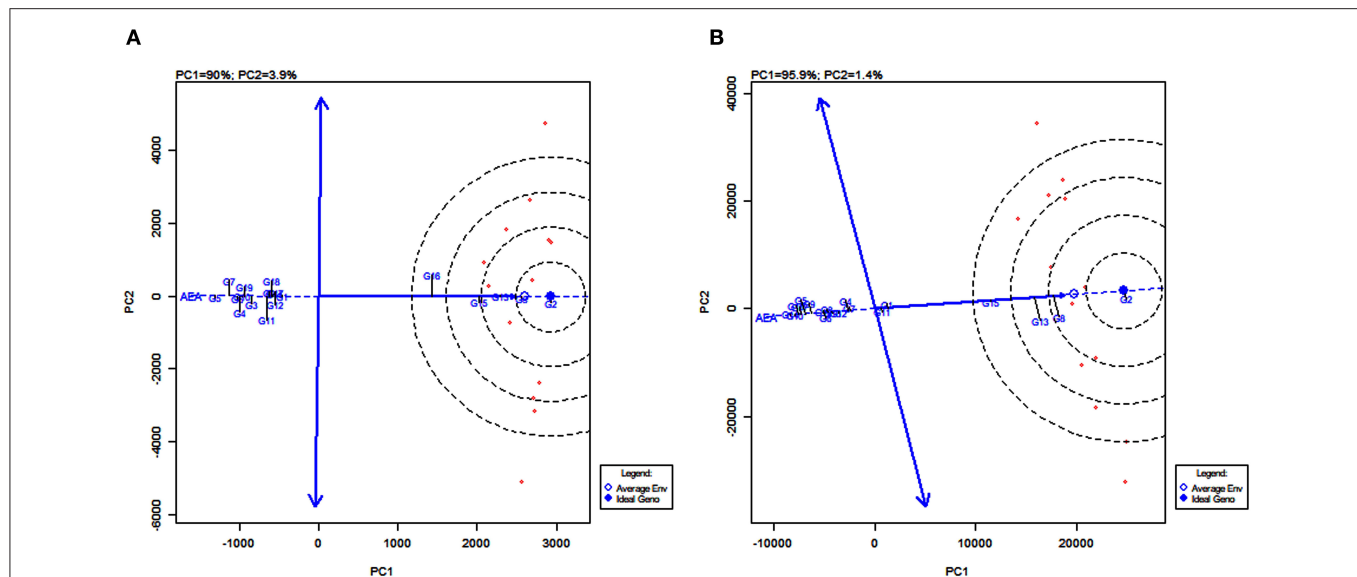


FIGURE 8 | GGE biplot of stability and mean performance of 19 sorghum varieties across average environments of 13 test environments for grain yield (A) and fodder yield (B).

DISCUSSION

In this study, a total of 108 landraces were collected, and 36 superior landraces of 108 landraces were identified. About 11 of these landraces were used to develop 19 sorghum varieties. Previous reports (Tesso et al., 2008; Adugna, 2014; Abraha et al., 2015; Amelework et al., 2016; Wondimu et al., 2020) have identified sorghum landraces with a wide range of variations that could provide new sources of tolerance and highly contrasting lines for use in future breeding programs. The landraces used in the present study were collected from the interior parts of central and southern India where they might have evolved in response to diverse agroecological zones and farming systems practiced in those regions and were better adapted to the low input and marginal cultivation conditions of these areas coupled with a frequent occurrence of grain mold disease. The 13 rainfed locations tested in this study represented the rainfed dryland conditions and were characterized by a complex climate that is largely semiarid and dry subhumid, with a short wet season followed by a long dry season. Sorghum cultivated in these regions is prone to highly erratic rainfall (spatially and temporally), with a strong risk of dry spells at critical growth stages and heavy rains at the grain maturity stage. The present study identified 36 superior performing landraces, cultivated in water-stress conditions and had a tolerance to grain mold due to an indirect selection for associated traits such as panicle shape, grain color, seed compactness, etc. over the years, which could be utilized to develop elite lines in sorghum breeding programs.

In the present study, the evaluation of 19 sorghum varieties derived from the superior landraces using AMMI and GGE biplots has demonstrated a higher contribution by the $G \times E$ interaction to the total variation than the genotypes suggesting that the environment had a high impact on the performance of

the sorghum varieties for both grain yield and fodder yield. Even though the proportion of the environment is the largest for both grain and fodder yields, the genotypes and $G \times E$ interaction have a paramount importance for the genotype evaluation (Yan and Kang, 2003). These findings are in agreement with Reddy et al. (2014) and Al-Naggar et al. (2018) who reported the predominance of the environment's main effect as the source of variation in the multi-environment trials in sweet sorghum and grain sorghum, respectively. Abiotic factors such as soil moisture, pH, mineral availability along with the weather and biotic factors including natural pest and disease occurrence might have contributed to large variations in the yield performance of the genotypes. The potential of the genotypes could be more exploited if the best performing genotypes were identified for the specific environments.

The total variations (56.6, 66.8%) contributed by IPCA1 and IPCA2 explained the differential performance of genotypes for grain yield and fodder yield across the 13 environments. Because of their maximum contribution, IPCA1 and IPCA2 were used to plot a two-dimensional GGE biplot. Gauch (2013) suggested that the most accurate model for AMMI can be predicted by using the first two IPCAs. Several researchers used the first two IPCAs for the GGE biplot analysis because they explained a greater percentage of the $G \times E$ interaction for sorghum (Al-Naggar et al., 2018), barley (Vaezi et al., 2017), pigeonpea (Rao et al., 2020; Kumar et al., 2021), wheat (Verma et al., 2015), and maize (Solomon et al., 2008).

AMMI stability value, which is the quantitative stability value developed through the AMMI model by Purchase et al. (2000), has been considered as the most appropriate single method of describing the stability of genotypes (Naroui et al., 2020). Several studies have identified the genotypes with smaller ASV and better stability and those with high ASV but higher yields for specific

TABLE 7 | Disease reaction scores against grain mold, IPCA scores, and ASV values of 19 sorghum genotypes over three environments from 2016 to 2018.

S. No.	Genotype/Environment	Score	PC1	PC2	ASV
1	PYPS-1	4.52	−0.07	0.08	0.01
2	PYPS-2	3.92	0.24	0.06	0.60
3	PYPS-3	4.50	0.02	0.11	0.11
4	PYPS-4	4.10	−0.16	−0.11	0.40
5	PYPS-5	4.51	0.03	0.07	0.09
6	PYPS-6	4.73	−0.13	0.03	0.32
7	PYPS-7	4.18	−0.07	0.06	0.18
8	PYPS-8	3.71	0.01	−0.04	0.04
9	PYPS-9	4.78	0.00	−0.09	0.09
10	PYPS-10	5.10	−0.13	0.14	0.37
11	PYPS-11	3.65	0.09	0.07	0.23
12	PYPS-12	4.80	−0.07	0.11	0.06
13	PYPS-13	3.97	0.00	−0.09	0.09
14	PYPS-14	5.32	0.03	0.02	0.07
15	PYPS-15	4.54	−0.03	−0.06	0.08
16	PYPS-16	5.00	0.01	−0.02	0.02
17	PYPS-17	4.74	0.20	−0.03	0.50
18	PYPS-18	4.73	−0.12	−0.07	0.72
19	PYPS-19	4.39	0.09	−0.04	0.32
20	SPV 462 (Susceptible check)	8.33	0.02	−0.13	0.24
21	IS 8545 (Resistant check)	4.04	0.03	−0.03	0.06
	E1	4.66	0.30	−0.17	0.33
	E2	4.63	0.04	0.29	0.29
	E3	4.64	−0.34	−0.12	0.85
	LSD (0.05)	2.54			

IPCA, Interaction principal component axis; ASV, AMMI stability value; LSD, Least significant difference.

Disease Score between 3.0 and 5.0 with 6–30% grain mold incidence is considered as tolerant, disease score 6.0–9.0 with 31–100% grain mold incidence and identified as susceptible.

adaptability in crops including bread wheat (Farshadfar, 2008), grain sorghum (Adugna, 2014), and finger millet (Lule et al., 2014). In the present study, the high yields and stability (low ASV) of varieties such as PYPS 2, PYPS 5, PYPS 8, and PYPS 13 might be attributed to the 11 parental landraces, which could offer potential new sources of genes for higher grain and fodder yields and stability. These genotypes merit further genetic studies for adaptation and physiological traits for dissecting the traits contributing to stability.

In this study, several stable sorghum varieties were identified for grain yield and fodder yield. For example, PYPS 8 was stable and high in yield for grain yield. However, most of the stable genotypes need not necessarily have the best yield performance (Mohammadi and Amri, 2007). Hence, GSI was used as a single criterion to classify stable genotypes. To our knowledge, this is the first study incorporating grain and fodder yields' mean, protein content, grain mold reaction, and stability index to identify high-yielding, high protein, grain mold tolerant, and stable sorghum varieties such as PYPS 1, PYPS 2, PYPS 8, PYPS

13, etc. In wheat, Singh et al. (2019) have used GSI to identify stable high-yielding genotypes in India.

Genotypic stability is crucial to grain yield and the best genotype needs to combine good yield and stable performance across a range of production environments. Based on the AMMI1 biplot, the current study identified the sorghum varieties with wide and specific adaptability for both grain yield and fodder yield. For example, sorghum variety PYPS 8 with a low absolute IPCA1 score and the above-average grain yield was stable, showing less-variable yield across the environments, thus making it a promising multilocation testing and validation. On the contrary, the varieties PYPS 1 and PYPS 10 performed consistently across locations but with below-average grain yields. The variety PYPS 2 showed specific adaptability to E1 and E10 for grain yield. Similarly, Al-Naggar et al. (2018) identified two grain sorghum B-lines with site-specific adaptability in Egypt.

Interestingly, the varieties differed in their grain yield and fodder yield performances across locations. For example, the variety PYPS 15 showed specific adaptation to environments E1 and E10 for high grain yield (3,432 kg/ha) but was the most widely adaptable variety for fodder yield (16,955 kg/ha). Sorghum varieties PYPS 2, PYPS 8, PYPS 13, and PYPS 15 were considered as the best dual purpose cultivars (Figure 9) due to their higher grain and fodder yields (3,698, 2,0586 kg/ha; 3,584, 18,632 kg/ha; 3,514, 18,122 kg/ha; 3,432, 16,955 kg/ha) whereas PYPS 16 and PYPS 11 were the best varieties exclusively for grain (3,268 kg/ha) and fodder (13,895 kg/ha).

The relationship between the testing environments was graphically evaluated by using the angles between the vectors. The presence of wide obtuse angles between E5 and E12 with E2 suggested a negative correlation or strong crossover $G \times E$ interaction for grain yield (Yan and Tinker, 2006). This indicated that the genotypes performing better in one environment would perform poorly in another environment. Closer relationships among the locations depicted by small cosine angles ($<90^\circ$) indicated the non-existence of a crossover $G \times E$ interaction suggesting that the ranking of genotypes does not change from environment to environment. The environments E5 and E12 and the environments E3 and E7 fall under the latter. The present study indicated a mixture of crossover and non-crossover types of $G \times E$ interaction, which has been reported in various studies (Rakshit et al., 2012; Naroui et al., 2013; Aruna et al., 2015). Furthermore, eliminating similar environments from multilocation trials of sorghum will help in the optimal utilization of resources.

In the polygon view of the GGE biplot derived from the first two main PC, PC1 refers to the yield ratio, associated with genotypic characteristics and PC2 refers to the yield related to the $G \times E$ interaction (Yan et al., 2007; Yan and Holland, 2010). In the present study, the contribution of the first two PCs toward 99.9 and 95.9% of the variability for grain yield and fodder yield, respectively, justified the use of the GGE biplot to effectively interpret the variability in the multi-environment data. The 13 tested environments in this study contributed to 35.5% of the total variation in grain yield and 28.9% in fodder yield. Mushayi et al. (2020) reported as much as 63% of the variation being explained by location in grain yield for maize. In this study, the

TABLE 8 | Nutrient composition of sorghum genotypes collected from E1 (Palem, Telangana, India) in 2018.

S. No.	Genotype	Starch	Sugar	Fe (mg/100 g)	Zn (mg/100 g)	Protein (%)
1	PYPS-1	49.73	8.76	5.60	3.40	12.32
2	PYPS-2	53.93	10.01	10.70	2.80	14.52
3	PYPS-3	38.44	12.68	5.50	2.40	9.78
4	PYPS-4	48.61	14.09	12.70	2.90	10.30
5	PYPS-5	53.88	11.43	5.40	1.90	10.52
6	PYPS-6	43.50	9.13	6.20	2.20	9.63
7	PYPS-7	45.13	12.18	7.40	1.90	9.25
8	PYPS-8	48.77	9.67	6.60	3.40	13.26
9	PYPS-9	52.86	5.25	6.00	3.30	14.13
10	PYPS-10	39.46	12.30	7.80	2.50	9.58
11	PYPS-11	45.43	9.84	5.80	2.70	10.73
12	PYPS-12	46.50	9.95	5.30	2.10	10.95
13	PYPS-13	57.09	8.82	4.90	2.40	10.66
14	PYPS-14	43.05	6.77	5.40	3.00	10.52
15	PYPS-15	40.16	14.93	4.40	1.80	8.78
16	PYPS-16	32.11	13.59	6.40	2.90	10.77
17	PYPS-17	56.38	6.98	5.40	2.00	11.28
18	PYPS-18	38.29	9.77	4.40	2.30	10.60
19	PYPS-19	45.20	9.60	6.70	2.80	10.72
	LSD	7.21	2.84	2.14	0.64	2.15
	CV%	14.71	11.41	8.75	7.86	14.24
	Range	24.98	9.68	8.30	1.6	5.74

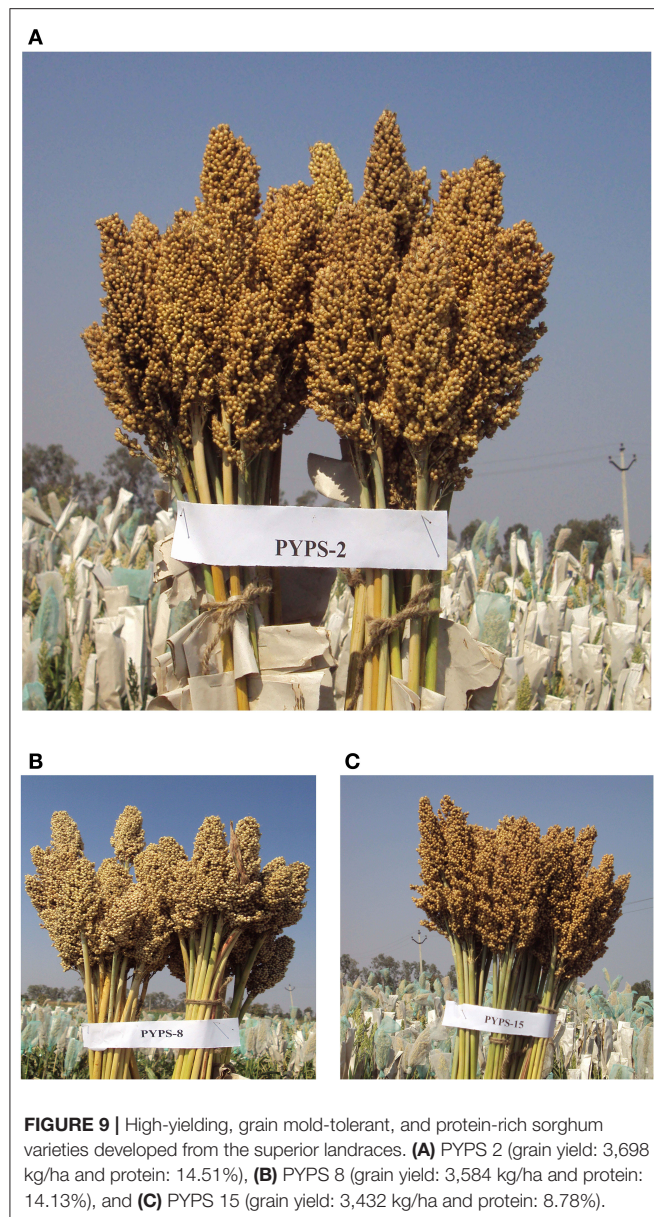
LSD, Least significant difference; CV, Coefficient of variation.

highest grain yielding variety PYPS 2 performed best at E2, E1, E11, E6, and E10. These environments are in complete contrast to the best yielding environments E3, E8, E12, and E13 for high fodder yield. Several authors identified high-yielding genotypes for grain sorghum (Rakshit et al., 2012; Al-Naggar et al., 2018), forage sorghum (Aruna et al., 2015), and sweet sorghum (Rono et al., 2016).

In the GGE biplot analysis, partitioning of the test locations into meaningful mega-environments is the best approach to exploit the positive $G \times E$ interaction (Yan and Tinker, 2006). In the present study, the test locations were divided into a total of four mega-environments for grain yield and four for fodder yield. The varieties performed differently across the mega-environments for grain yield and fodder yield. For example, the variety PYPS 2 was the best for the two mega-environments for high grain yield and fodder yield and PYPS 5 was not suitable for any mega-environment suggesting that different genotypes should be deployed for each mega-environment to achieve optimal adaptation.

Based on the AMMI analysis, PYPS 8 and PYPS 13 were identified as the most stable sorghum varieties with a low grain mold incidence. The AMMI and GGE biplot methods were used to identify resistance sources to different diseases in multiple locations and also high-yielding stable genotypes in wheat, groundnut, mungbean, melon, etc. (Akcura et al., 2017; Chaudhari et al., 2019; Naroui et al., 2020; Tollo et al., 2020). In a previous study, Diatta et al. (2019) evaluated five sorghum

hybrids along with their parental lines and reported a significant $G \times E$ interaction for grain yield. However, the $G \times E$ interaction was not significant for panicle mold infestation. Grain mold is a complex disease whose incidence is governed by a host resistance and an environment. Grain mold resistance was correlated with open panicles, long glumes, a greater glume coverage length, and area (Sharma et al., 2010). Loose panicle sorghum is more likely to be resistant to grain mold because the compact heads hold more moisture that favors disease development. Weather factors, particularly relative humidity play a dominant and determining role in grain mold severity. The wet weather condition following flowering is necessary for the grain mold development. The longer the duration of wetness on grain surface, the greater is the incidence of grain mold (Das et al., 2020). Photoperiod-sensitive cultivars that mature after the rains often escape grain mold infection (Patted et al., 2011). Sorghum cultivars with a white grain are more vulnerable to grain mold than those with a brown and red grain pericarp. In the present study, the varieties PYPS 8 and PYPS 13, which showed a low grain mold incidence, had a brown and yellow pericarp, respectively. Their parental lines viz., PSLRC 8, PSLCRC 9, PSLRC 2, and PSLRC 21 were characterized by a brown and yellow pericarp with high tannin levels, which might have contributed to grain mold tolerance. Furthermore, their panicles were semi-compact to loose a trait important for grain mold tolerance due to non-retention of moisture and better air circulation (Glueck et al., 1977; Rao et al., 1999; Patted et al., 2011). Glume length and area of coverage over the grain is related to the grain mold escape



as the grains are protected from the exposure to rain. Patted et al. (2011) reported that sorghum progenies with a very long-to-long glume coverage escaped grain mold tolerance. They also reported that black- and red-colored glumes were moderately resistant to grain mold due to the presence of tannins, which inhibit the growth of saprophytic fungi, and thus reducing a grain mold incidence.

An analysis of nutrient composition revealed a mean starch content of 46.04% that is lower than the average content varying from 56 to 73% (Ratnavathi and Patil, 2013). The protein content ranging from 10.4 to 10.62% was reported in sorghum germplasm collections (Weckwerth et al., 2020). Other studies in sorghum have reported the protein content of 9–11% (Elbashir et al., 2008; Ahmed et al., 2014). Abdelhalim et al. (2021) reported

the protein content ranging from 6.3 to 12.5% among the five landraces evaluated for their potential in biofortification. The protein content of the three varieties *viz.*, PYPS 2 (14.52%), PYPS 9 (14.13), and PYPS 8 (13.26%) tested in this study is higher than that of sorghum landraces reported by Abdelhalim et al. (2021). The high protein content of the genotypes in this study suggested the possibility of a similar feature in the parent landraces PSLRC 2, PSLRC 3, PSLRC 4, PSLRC 8, PSLRC 9, and PSLRC 10. These landraces, in addition to the genotypes PYPS 2, PYPS 8, and PYPS 9, offer a significant source of new genetic materials, as well as fortified ingredients for enhancing the nutritional value of sorghum grains. One nutritional constraint to the use of sorghum as food is the poor digestibility of sorghum proteins in cooking (Tesso et al., 2008). Duressa et al. (2020) have reported sorghum accessions with a protein content as high as 15% but with different digestibility percentages (a measure of the susceptibility of protein to proteolysis). For example, the sorghum accession SC663 with a protein content of 15.5% had a high protein digestibility of 70.77%. On the contrary, the sorghum accession SC25 with a protein content of 14.8% had a protein digestibility of 49.73%. Hence, it is important that the sorghum varieties with a high protein content are evaluated for protein digestibility to identify the genotypes with high protein availability.

In the present study, the varieties varied in their grain Fe and Zn concentrations. Phuke et al. (2017) reported a highly significant $G \times E$ interaction for Fe and Zn levels in sorghum recombinant inbred lines. They found a significant variation in Fe and Zn across the environments and a significant positive correlation between the Fe and Zn concentrations. While the varied concentrations of Fe and Zn in the present study might be genetically controlled, the effect of the environment and the possible $G \times E$ interaction need to be investigated by multi-environment testing of the sorghum genotypes. Shegro et al. (2012) also reported a diversity in nutritional composition among the sorghum landraces and identified the accessions with high concentrations of proteins and minerals (Fe, Fe, Ca, K, Mn, P, and Mg). The variations in mineral contents of the 19 sorghum genotypes in this study may be due to genotype, mineral concentrations in the soil as well as translocation rates of the elements from the soil, as well as environmental factors such as temperature and rainfall or an inherent ability of the genotypes to absorb the nutrients from the soil (Shegro et al., 2012). The assessment of a phenotypic correlation among the protein, starch, sugar, and mineral compositions of the genotypes in this study might have implications for the sorghum crop improvement in relation to human nutrition and livestock feed. Weckwerth et al. (2020) have suggested the use of genomic selection (GS) using genome-wide DNA markers in crop breeding programs to target multiple and complex traits. This can be deployed in a sorghum breeding program to simultaneously target yield, drought tolerance, and nutritional composition including protein, starch, Fe, and Zn. Not just for sorghum, but such an approach integrating GS with an environment-dependent PANOMICS analysis can improve the productivity, biotic and abiotic stress tolerance, and nutritional value of several crops including pigeonpea, groundnut, chickpea, etc. (Weckwerth et al., 2020).

In cereals, the nutritional quality and end use properties are linked to arabinoxylans (AX) that are the major polymers present in the cell walls of grains (Saulnier et al., 2007). Though wheat AX has been extensively reviewed (Saulnier et al., 2007), limited information is available on sorghum AX concerning the amount, composition, and expression. Nandini and Salimath (2001) reported that the structural features of AX are linked with a good flatbread-making quality in sorghum. PYPS-2, the high yielding, protein-rich, and grain mold-tolerant genotype has a high consumer preference owing to good flatbread-making quality (Jaisimha, 2019), which might be linked to the structure and/or amount of AX. Though the expression is mainly genetically controlled, a positive relationship with the amount of AX and drought has been reported in wheat (Coles et al., 1997). As demonstrated in this study, the 36 landraces and the 19 genotypes were evaluated under rainfed situations where prolonged drought stress is a common occurrence, which may have influenced the AX expression *vis-à-vis* flat bread-making quality. The evaluation of sorghum genotypes for their AX amount and expression in different environments might help in identifying stable and adaptable cultivars with enhanced nutritional quality. To achieve this, research may first be reoriented toward the understanding of the AX biosynthesis, the identification of the candidate genes, and exploitation of the variation toward the development of new cultivars with improved nutritional quality and processing properties.

CONCLUSION

The present study reports the development of an elite sorghum breeding material using locally adapted landraces through pedigree breeding for the grain and fodder yield. The study has identified stable environments for grain yield (E5) and fodder yield (E11), which suggest that testing of initial hybrids and varietal trials in these two environments is more discriminating and rewarding to save resources in Telangana, India. The study showed the potential of the collected landraces for the development of high-yielding sorghum varieties suitable for rainfed cultivation in India and spill-over to African subcontinents. The study demonstrated a scope for the utilization of these diverse, locally adapted landraces that have been evolved

through a continuous selection by the farmers in varied-agro-climatic zones under low input conditions to develop climate-resilient sorghum cultivars and ultimately contribute to healthy, global food, and feed security.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MN, VR, MG, and MP designed the experiment. MN and VR carried out the experiments. MN, VR, MG, and SG analyzed the data. CS, SM, SG, and HK conducted the field experiments with logistical support by MV, KA, and RJ. MS helped in nutrient analysis. MN, VR, and MG wrote the draft manuscript. All authors made contributions toward the compilation of the draft and approved the final manuscript.

FUNDING

This study was funded by National Innovation on Climate Resilient Agriculture, ICAR-Central Research Institute for Dryland Agriculture, Hyderabad, India under the project Risk minimization in drought-prone Telangana districts through millet-based crop diversification.

ACKNOWLEDGMENTS

The authors are grateful to Prof. Jayashankar Telangana State Agricultural University, Rajendranagar, Hyderabad, India for providing institutional support throughout the implementation of the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.659874/full#supplementary-material>

REFERENCES

- Abdelhalim, T. S., Abdelhalim, N. S., Kamal, N. M., Mohamed, E. E., and Hassan, A. B. (2021). Exploiting the potential of Sudanese sorghum landraces in biofortification: physicochemical quality of grain sorghum (*Sorghum bicolor* L. Moench) landraces. *Food Chem.* 337:127604. doi: 10.1016/j.foodchem.2020.127604
- Abdelhalim, T. S., Kamal, N. M., and Hassan, A. B. (2019). Nutritional potential of wild sorghum: grain quality of Sudanese wild sorghum genotypes (*Sorghum bicolor* L. Moench). *Food Sci. Nutr.* 7, 1529–1539. doi: 10.1002/fsn3.1002
- Abdi, H., and Williams, L. J. (2010). Principal component analysis. wiley interdisciplinary reviews. *Comput. Stat.* 2, 433–459. doi: 10.1002/wics.101
- Abraha, T., Githiri, S. M., Kasili, R., Araia, W., and Nyende, A. B. (2015). Genetic variation among sorghum (*Sorghum bicolor* [L.] Moench) landraces from Eritrea under post-flowering drought stress conditions. *Am. J. Plant Sci.* 6, 1410–1424. doi: 10.4236/ajps.2015.69141
- Admas, S., and Tesfaye, K. (2017). Genotype-by-environment interaction and yield stability analysis in sorghum (*Sorghumbicolor* (L.) Moench) genotypes in North Shewa, Ethiopia. *Acta Univ. Sapient. Agric.* 9, 82–84. doi: 10.1515/ausae-2017-0008
- Adujna, A. (2014). Analysis of *in situ* diversity and population structure in Ethiopian cultivated *Sorghum bicolor* (L.) landraces using phenotypic traits and SSR markers. *Springer Plus.* 3, 1–14. doi: 10.1186/2193-1801-3-212
- Ahmed, S. O., Abdalla, A. W. H., Inoue, T., Ping, A., and Babiker, E. E. (2014). Nutritional quality of grains of sorghum cultivar grown under different levels of micronutrients fertilization. *Food Chem.* 159, 374–380. doi: 10.1016/j.foodchem.2014.03.033

- Akcura, M., Akan, K., and Hocaoglu, O. (2017). Biplot analysis of leaf rust resistance in pure lines selected from eastern Anatolian bread wheat landraces of Turkey. *Turk. J. Field Crops*. 22, 227–234. doi: 10.17557/tjfc.356231
- Akinseye, F. M., Ajeigbe, H. A., Traore, P. C. S., Agele, S. O., Zemadec, B., and Whitbread, A. (2020). Improving sorghum productivity under changing climatic conditions: a modelling approach. *Field Crops Res.* 246, 107685–107696. doi: 10.1016/j.fcr.2019.107685
- Al-Naggar, A. M., El-Salam, R. M., Asran, M. R., and Yaseen, W. (2018). Yield adaptability and stability of grain sorghum genotypes across different environments in Egypt using AMMI and GGE-biplot Models. *Annu. Res. Rev. Biol.* 23, 1–16. doi: 10.9734/ARRB/2018/39491
- Amelework, B., Shimelis, H., Tongoona, P., Laing, M., and Mengistu, F. (2015). Genetic variation in lowland sorghum (*Sorghum bicolor* (L.) Moench) landraces assessed by simple sequence repeats. *Plant Genet. Resour.* 13, 131–141. doi: 10.1017/S1479262114000744
- Amelework, B., Shimelis, H., Tongoona, P., Laing, M., and Mengistu, F. (2016). Genetic diversity of lowland sorghum landraces assessed by morphological and microsatellite markers. *Aust. J. Crop Sci.* 10, 291–298. doi: 10.21475/ajcs.2016.10.03.p6665
- Ananda, G. K. S., Myrans, H., Norton, S. L., Gleadow, R., Furtado, A., and Henry, R. J. (2020). Wild sorghum as a promising resource for crop improvement. *Front. Plant Sci.* 11:1108. doi: 10.3389/fpls.2020.01108
- AOAC. (2016). *Official Methods of Analysis of the Association of Analytical Chemists, 20th Edn.* Rockville, MD: AOAC International.
- Aruna, C., Rakshit, S., Shrotria, P., Pahuja, S., Jain, S., Kumar, S., et al. (2015). Assessing genotype-by-environment interactions and trait associations in forage sorghum using GGE biplot analysis. *J. Agric. Sci.* 154, 73–86. doi: 10.1017/S0021859615000106
- Awika, J. M. (2017). Sorghum: its unique nutritional and health-promoting attributes. *Gluten Free Ancient Grains* 2017, 21–54. doi: 10.1016/B978-0-08-100866-9.00003-0
- Balcerek, P. A., Frankowski, J., and Stuper-Szablewska, K. (2020). The influence of weather conditions on bioactive compound content in sorghum grain. *Eur. Food Res. Technol.* 246, 13–22. doi: 10.1007/s00217-019-03391-0
- Bavandpori, F., Ahmad, J., and Hossaini, S. (2018). Stability analysis of bread wheat landraces and lines using biometrical genetic models. *Genetika*. 50, 449–464. doi: 10.2298/GENSRI1802449B
- Brar, D. S., and Khush, G. S. (2018). “Wild relatives of rice: A valuable genetic resource for genomics and breeding research,” in *Compendium of Plant Genomes: The Wild Oryza Genomes*, eds T. K. Mondal and R. J. Henry (Berlin: Springer-Verlag), 1–25. doi: 10.1007/978-3-319-71997-9_1
- Chaudhari, S., Khare, D., Patil, S. C., Sundravada, S., Variath, M. T., Sudini, H. K., et al. (2019). Genotype \times environment studies on resistance to late leaf spot and rust in genomic selection training population of peanut (*Arachis hypogaea* L.). *Front. Plant Sci.* 10:1338. doi: 10.3389/fpls.2019.01338
- Coles, G. D., Hartunian-Sowa, S. M., Jamieson, P. D., Hay, A. J., Atwell, W. A., and Fulcher, R. G. (1997). Environmentally-induced variation in starch and non-starch polysaccharide content in wheat. *J. Cereal Sci.* 26, 47–54. doi: 10.1006/jcrs.1996.0102
- Cuevas, H. E., Prom, L. K., Cooper, E. A., Knoll, J. E., and Ni, X. (2018). Genome-wide association of mapping of anthracnose (*Colletotrichum sublineolum*) resistance in the U.S. Sorghum Association Panel. *Plant Genome* 11:170099. doi: 10.3835/plantgenome2017.11.0099
- Das, I. K., Aruna, C., and Tonapi, V. A. (2020). *Sorghum Grain Mold*. Hyderabad: Indian Council of Agricultural Research –Indian Institute of Millets Research.
- Davis, K. F., Chhatre, A., Rao, N. D., Singh, D., and De Fries, R. (2019). Sensitivity of grain yields to historical climate variability in India. *Environ. Res. Lett.* 14:064013. doi: 10.1088/1748-9326/ab22db
- Dehghani, H., Sabaghpour, S. H., and Ebadi, A. (2010). Study of Genotype \times Environment interaction for chickpea yield in Iran. *Agron. J.* 102, 1–8. doi: 10.2134/agronj2009.0156
- Dereese, S. A., Shimelis, H., Mwandzingeni, L., and Liang, M. (2018). Agromorphological characterization and selection of sorghum landraces. *Acta Agric. Scand.* 68, 585–595. doi: 10.1080/09064710.2018.1448884
- Diatla, C., Tovignan, T., Adoukonou, S. H., Aidara, O., Diao, Y., Sarr, M., et al. (2019). Development of sorghum hybrids for stable yield and resistance to grain mold for the Center and South-East of Senegal. *Crop Prot.* 119, 197–207. doi: 10.1016/j.cropro.2019.02.001
- Duessa, D., Bean, S., Amand, P. S., and Tesso, T. (2020). Identification of variant α -kafirin alleles associated with protein digestibility in grain sorghum (*Sorghum bicolor* (L.) Moench). *Crop Sci.* 60, 2467–2478. doi: 10.1002/csc.20198
- Dwivedi, S. L., Ceccarelli, S., Blair, M. W., Upadhyaya, H. D., Are, A. K., and Ortiz, R. (2016). Landrace germplasm for improving yield and abiotic stress adaptation. *Trends Plant Sci.* 21, 31–42. doi: 10.1016/j.tplants.2015.10.012
- Dykes, L., Peterson, G. C., Rooney, W. L., and Rooney, L. W. (2011). Flavonoid composition of lemon yellow sorghum genotypes. *Food Chem.* 128, 173–179. doi: 10.1016/j.foodchem.2011.03.020
- Eggen, M., Ozdogan, M., Zaitchik, B., Ademe, D., Fotz, J., and Simane, B. (2019). Vulnerability of sorghum production to extreme, sub-seasonal weather under climate change. *Environ. Res. Lett.* 14:045005. doi: 10.1088/1748-9326/aaf19
- Elangovan, M., Ganesamurthy, K., Rajaram, S., Sankarapandian, K., and Kiranbabu, P. (2012). Collection and conservation of sorghum landraces from Tamil Nadu. *Electronic J. Plant Breed.* 3, 753–762.
- Elangovan, N., Prabhakar, T. V. A., and Chandra Sekhara Reddy, D. (2009). Collection and characterization of Indian sorghum landraces. *Indian J. Plant Genet. Resour.* 22, 173–181.
- Elbashir, H. Z., Mustafa, A. I., El Tinay, A. H., and Babiker, E. E. (2008). Biochemical characteristics of sorghum (*Sorghum bicolor* L. Moench) flour supplemented with cluster bean (*Cyamopsis tetragonoloba* L.): Influence of fermentation and/or cooking. *J. Biol. Sci.* 8, 722–729. doi: 10.3923/jbs.2008.722.729
- FAOSTAT. (2019). *Food and Agriculture Organization of the United Nations*. Rome: FAOSTAT.
- Farshadfar, E. (2008). Incorporation of AMMI stability value and grain yield in a single non parametric index (Genotype Selection Index) in bread wheat. *Pakistan J. Biol. Sci.* 11, 1791–1796. doi: 10.3923/pjbs.2008.1791.1796
- Gauch, H. G. (2013). A simple protocol for AMMI analysis of yield trials. *Crop Sci.* 53, 1860–1869. doi: 10.2135/cropsci2013.04.0241
- Gauch, H. G., and Zobel, R. (1996). “AMMI analysis of yield trials,” in *Genotype-by-Environment Interaction*, eds M. S. Kang and H. S. Gauch (Boca Raton, FL: CRC CRC), 85–122. doi: 10.1201/9781420049374.ch4
- Ghebru, B., Schmidt, R., and Bennetzen, J. (2002). Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. *Theor. Appl. Genet.* 105, 229–236. doi: 10.1007/s00122-002-0929-x
- Glueck, J. A., Rooney, L. W., Rosenow, D. T., and Miller, F. R. (1977). *Physical and Structural Properties of Field Deteriorated Sorghum Grain*. College Station, TX: Texas Agricultural Experiment Station.
- Godwin, I. D., Rutkoski, J., Varshney, R. K., and Hickey, L. T. (2019). Technological perspectives for plant breeding. *Theor. Appl. Genet.* 132, 555–557. doi: 10.1007/s00122-019-03321-4
- Goedhart, P. W., and Thissen, J. T. N. M. (2010). *Biometrics GenStat Procedure Library Manual, 13th Edn.* Wageningen: Wageningen Universiteit.
- Guerra, E., De Oliveira, R., Daros, E., Zambon, J., Ido, O., and Bespalhok, J. (2009). Stability and adaptability of early maturing sugarcane clones by AMMI analysis. *Crop Breed. Appl. Biotechnol.* 9, 260–267. doi: 10.12702/1984-7033.v09n03a08
- Hadebe, S. T., Modi, A. T., and Mabhaudhi, T. (2017). Drought tolerance and water use of cereal crops: a focus on sorghum as a food security crop in Sub-Saharan Africa. *J. Agron. Crop Sci.* 203, 177–191. doi: 10.1111/jac.12191
- Jaisimha, R. K. (2019). *Performance of Yellow Pericarp Sorghum Genotypes to Moisture Conservation Through Land Configuration and Vegetative Mulch During Kharif*. M. Sc. (Agril.) (Thesis). Prof. Jayashankar Tel. State Agric. Univ., Hyderabad, India.
- Kalaria, R. K., Patel, A., and Desai, H. (2020). Isolation and characterization of dominant species associated as grain mold complex of sorghum under south Gujarat region of India. *Indian Phytopathol.* 73, 159–164. doi: 10.1007/s42360-020-00196-0
- Kumar, A. A., Anuradha, K., Ramaiah, B., Grando, S., Frederick, H., Rattunde, W., et al. (2015). Recent advances in sorghum biofortification research. *Plant Breed. Rev.* 39, 89–124. doi: 10.1002/9781119107743.ch3
- Kumar, M. V. N., Ramya, V., Sameer Kumar, C. V., Raju, T., Sunil Kumar, N. M., Seshu, G., et al. (2021). Identification of pigeonpea genotypes with wider adaptability to rainfed environments through AMMI and GGE biplot analyses. *Indian J. Genet.* 81, 63–73. doi: 10.31742/IJGPB.81.1.7
- Kyrtatzis, A. C., Nikoloudakis, N., and Katsiotis, A. (2019). Genetic variability in landraces populations and the risk to lose genetic variation. The example of

- landrace 'Kyperounda' and its implications for *ex situ* conservation. *PLoS ONE* 11:e0224255. doi: 10.1371/journal.pone.0224255
- Lule, D., Fetene, M., DeVilliers, S., and Tesfaye, K. (2014). Additive main effects and multiplicative interactions (AMMI) and genotype by environment interaction (GGE) biplot analyses aid selection of high yielding and adapted finger millet varieties. *J. Appl. Biosci.* 76, 6291–6303. doi: 10.4314/jab.v76i1.1
- McCarthy, D. S., and Vlek, P. L. G. (2012). Impact of climate change on sorghum production under different nutrient and crop residue management in semi-arid region of Ghana: a modeling perspective. *African Crop Sci.* 20, 243–259.
- Mofokeng, M., Shimelis, H., and Laing, M. D. (2017). Agro-morphological diversity of South African sorghum genotypes assessed through quantitative and qualitative phenotypic traits. *S. Afr. J. Plant Soil.* 34, 1–10. doi: 10.1080/02571862.2017.1319504
- Mofokeng, M. A., Shimelis, H., Tongoona, P., and Laing, M. D. (2018). Protein content and amino acid composition among selected South African sorghum genotypes. *J. Food Chem. Nutr.* 6, 1–12. doi: 10.33687/jfcn.006.01.1927
- Mohammadi, R., and Amri, A. (2007). Comparison of parametric and non-parametric parameters for selecting stable and adapted durum wheat genotypes in variable environments. *Euphytica* 159, 419–432. doi: 10.1007/s10681-007-9600-6
- Mundia, C. W., Secchi, S., Akamani, K., and Wang, G. (2019). A regional comparison of factors affecting global sorghum production: the case of North America, Asia and Africa's Sahel. *Sustainability* 11:2135. doi: 10.3390/su11072135
- Mushayi, M., Shimelis, H., Derera, J., Shayanowako, A. L. T., and Mathew, I. (2020). Multi-environment evaluation of maize hybrids developed from tropical and temperate lines. *Euphytica* 216:84. doi: 10.1007/s10681-020-02618-6
- Nandini, C. D., and Salimath, P. V. (2001). Structural features of arabinoxylans from sorghum having good roti-making quality. *Food Chem.* 74, 417–422. doi: 10.1016/S0308-8146(01)00266-7
- Naroui, R., Mohammad, R., Abdul Kadir, M., Rafii Hawa, M. Y., Jaafar Naghavi, M. R., and Farzaneh, A. (2013). Genotype x environment interaction by AMMI and GGE biplot analysis in three consecutive generations of wheat (*Triticum aestivum*) under normal and drought stress conditions. *Aust. J. Crop. Sci.* 7, 956–961.
- Naroui, R., Mohammad, R., and Bakhshi, B. (2020). GGE biplot tool to identify melon fruit weight stability under different drought stress conditions. *Int. J. Veg. Sci.* 27, 220–223. doi: 10.1080/19315260.2020.1805538
- Nida, H., Girma, G., Mekonen, M., Lee, S., Seyoum, A., Dessalgen, K., et al. (2019). Identification of sorghum grain mold resistance loci through genome wide association mapping. *J. Cereal Sci.* 85, 295–304. doi: 10.1016/j.jcs.2018.12.016
- Ocheing, G., Nguji, K., Wamalwa, L. N., Manyasa, E., Muchira, N., Nyamongo, D., et al. (2020). Novel sources of drought tolerance from landraces and wild sorghum relatives. *Crop Sci.* 61, 104–118. doi: 10.1002/csc2.20300
- Patil, J. V., Reddy, P. S., Prabhakar, U. A. V., Sunil, G., and Ganapathy, K. N. (2014). History of post-rainy season sorghum research in India and strategies for breaking the yield plateau. *Indian J. Genet.* 74, 271–285. doi: 10.5958/0975-6906.2014.00845.1
- Patted, V., Kamatar, M. Y., and Deshpande, S. K. (2011). Variability for grain mold resistance in sorghum F₃ progenies of red x red and red x white grain crosses. *Elec. J. Plant Br.* 2, 248–252.
- Philipo, M., Ndadikemi, P. A., and Mbega, E. R. (2020). Environmental and genotypes influence on seed iron and zinc levels of landraces and improved varieties of common bean (*Phaseolus vulgaris* L.) in Tanzania. *Ecol. Gen. Genom.* 15:100056. doi: 10.1016/j.egg.2020.100056
- Phuke, R., Anuradha, K., Radhika, K., Jabeen, F., Ghanta, A., Thatikunta, R., et al. (2017). Genetic variability, genotype x environment interaction, correlation, and GGE biplot analysis for grain iron and zinc concentration and other agronomic traits in RIL population of sorghum (*Sorghum bicolor* L. Moench). *Front. Plant Sci.* 8:712. doi: 10.3389/fpls.2017.00712
- Pouresmael, M., Kanouni, H., Hajjhasani, M., Astraki, H., Mirakhorli, A., Nasrollahi, M., et al. (2018). Stability of chickpea (*Cicer arietinum* L.) landraces in national plant gene bank of Iran for drylands. *J. Agric. Sci. Technol.* 20, 387–400.
- Purchase, J. L., Hatting, H., and Vandeventer, C. S. (2000). Genotype x environment interaction of winter wheat (*Triticum aestivum* L.) in South Africa: II. Stability analysis of yield performance. *S. Afr. J. Plant Soil.* 17, 101–107. doi: 10.1080/02571862.2000.10634878
- Rakshit, S., Ganapathy, K. N., Gomashe, S. S., Rathore, A., Ghorade, R. B., Nagesh Kumar, M. V., et al. (2012). GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica* 185, 465–479. doi: 10.1007/s10681-012-0648-6
- Rao, A. S., Prasad, R. K. E., Mengesh, M. H., and Reddy, V. G. (1999). Morphological diversity in sorghum germplasm from India. *Gen. Resour. Crop Evol.* 43, 559–567. doi: 10.1007/BF00138832
- Rao, P. J. M., Kishore, N. S., Sandeep, S., Neelima, G., Rao, P. M., Das, D. M., et al. (2020). Evaluation of performance and yield stability analysis based on AMMI and GGE-Biplot in promising pigeonpea [*Cajanus cajan* (L.) Millsbaugh] genotypes. *Legume Res.* doi: 10.18805/LR-4299. [Epub ahead of print].
- Ratnavathi, C. V., and Patil, J. V. (2013). Sorghum utilization as food. *J. Nutr. Food Sci.* 4, 247–255.
- Reddy, P. S., Reddy, B. V. S., and Rao, P. S. (2014). Genotype by sowing date interaction effects on sugar yield components in sweet sorghum (*Sorghum bicolor* L. Moench). *SABRAO J. Breed Genet.* 46, 241–255.
- Rono, J., Cheruiyot, E., Othira, J., Njuguna, V., Macharia, J., Owuoch, J., et al. (2016). Adaptability and stability study of selected sweet sorghum genotypes for ethanol production under different environments using AMMI analysis and GGE biplots. *Sci. World J.* 2016, 1–14. doi: 10.1155/2016/4060857
- Saulnier, L., Sado, P. E., Branlard, G., and Charmet, G. (2007). Wheat arabinoxylans: exploiting variation in amount and composition to develop enhanced varieties. *J. Cereal Sci.* 46, 261–281. doi: 10.1016/j.jcs.2007.06.014
- Sharma, R., Rao, V., Upadhyaya, H., Reddy, V., and Thakur, R. (2010). Resistance to grain mold and downy mildew in a mini-core collection of sorghum germplasm. *Plant Dis.* 94, 439–444. doi: 10.1094/PDIS-94-4-0439
- Shagro, A., Shargie, N. G., and Biljon, A. (2012). Diversity in starch, protein and mineral composition of sorghum landrace accessions from Ethiopia. *J. Crop Sci. Biotechnol.* 15, 275–280. doi: 10.1007/s12892-012-0008-z
- Singh, C., Gupta, A., Gupta, V., Kumar, P., Sendhil, R., Tyagi, B. S., et al. (2019). G x E interaction analysis of multi-environment wheat trials in India using AMMI and GGE biplot models. *Crop Breed. Appl. Biotechnol.* 19, 309–318. doi: 10.1590/1984-70332019v19n3a43
- Singh, S. D., and Bandyopadhyay, R. (2000). "Grain mold," in *Compendium of Sorghum Diseases*, eds R. A. Frederiksen and G. N. Odvody (MN: American Phytopathological Society), 38–40.
- Solomon, A., Nigussie, M., and Zelleke, H. (2008). Genotype-environment interaction and stability analysis for grain yield of maize (*Zea mays* L.) in Ethiopia. *Asian J. Plant Sci.* 7, 163–169. doi: 10.3923/ajps.2008.163.169
- Tesfamichael, A., Githiri, F., Kasili, R., Skilton, R., Maina, S., and Nyende, A. B. (2014). Genetic diversity analysis of eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm using SSR markers. *Mol. Plant Breed.* 5, 1–12. doi: 10.5376/mpb.2014.05.0013
- Tesso, T., Kapran, I., Grenier, C., Snow, A., Sweeney, P., Pedersen, J., et al. (2008). The potential for crop-to-wild gene flow in sorghum in Ethiopia and Niger: a geographic survey. *Crop Sci.* 48, 1425–1431. doi: 10.2135/cropsci2007.08.0441
- Thakur, R. P., Reddy, B. V. S., and Mathur, K. (2007). *Screening Techniques for Sorghum Diseases. Information Bulletin No. 76*. Hyderabad, Patancheru: International Crops Research Institute for Semi-Arid Tropics.
- Tollo, J. A., Ojwang, O., Pascal, P., Karimi, R., and Mafurah, J. J. (2020). Genotype-by-environment interaction and stability of resistance in mungbean landraces against common bacterial blight across semi-arid environments. *Euphytica* 216:175. doi: 10.1007/s10681-020-02705-8
- Vaezi, B., Pour-Aboughadareh, A., Mohammadi, R., Armion, M., Mehraban, A., Hossein-Pour, T., et al. (2017). GGE biplot and AMMI analysis of barley yield performance in Iran. *Cereal Res. Commun.* 45, 500–511. doi: 10.1556/0806.45.2017.019
- Vara Prasad, B. V., and Sridhar, V. (2019). Diversity studies in yellow pericarp sorghum (*Sorghum bicolor* (L.) Moench) genotypes for yield attributes. *Int. J. Curr. Microbiol. Appl. Sci.* 8, 361–366. doi: 10.20546/ijcm.2019.812.048
- Verma, A., Chatrath, R., and Sharma, I. (2015). AMMI and GGE biplots for G x E analysis of wheat genotypes under rain fed conditions in central zone of India. *J. Appl. Nat. Sci.* 7, 656–661. doi: 10.31018/jans.v7i2.662

- Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P., and Varshney, R. (2020). PANOMICS meets germplasm. *Plant Biotechnol.* 18, 1507–1525. doi: 10.1111/pbi.13372
- Wondimu, Z., Bantte, K., Paterson, A., and Worku, W. (2020). Agromorphological diversity of Ethiopian sorghum [*Sorghum bicolor* (L.) Moench] landraces under water limited environments. *Genet. Resour. Crop Evol.* 67, 2149–2160. doi: 10.1007/s10722-020-00968-7
- Yan, W., and Holland, J. B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica* 171, 351–369. doi: 10.1007/s10681-009-0030-5
- Yan, W., Hunt, L. A., Sheng, Q., and Szlavniks, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40, 597–605. doi: 10.2135/cropsci2000.403597x
- Yan, W., and Kang, M. (2003). *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists*. Boca Raton, FL: CRC Press. doi: 10.1201/9781420040371
- Yan, W., Kang, M., Ma, B., Woods, S., and Cornelius, P. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* 47:374. doi: 10.2135/cropsci2006.06.0374
- Yan, W., and Rajcan, I. (2002). Biplots analysis of the test sites and trait relations of soybean in Ontario. *Crop Sci.* 42, 11–20. doi: 10.2135/cropsci2002.0011
- Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi environment trial data: Principles and applications. *Canad. J. Plant Sci.* 86, 623–645. doi: 10.4141/P05-169

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Nagesh Kumar, Ramya, Govindaraj, Sameer Kumar, Maheshwaramma, Gokenpally, Prabhakar, Krishna, Sridhar, Venkata Ramana, Avil Kumar and Jagadeeshwar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Genotypic Variation in Cultivated and Wild Sorghum Genotypes in Response to *Striga hermonthica* Infestation

Nicoleta Muchira^{1,2}, Kahi Ngugi^{1*}, Lydia N. Wamalwa¹, Millicent Avosa², Wiliter Chepkorir², Eric Manyasa², Desterio Nyamongo³ and Damaris A. Odeny^{2*}

¹ Department of Plant Science and Crop Protection, University of Nairobi, Nairobi, Kenya, ² International Crops Research Institute for the Semi-Arid Tropics-Kenya, Nairobi, Kenya, ³ Kenya Agricultural and Livestock Research Organization, Genetic Resources Research Institute, Kikuyu, Kenya

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Therese Bengtsson,
Swedish University of Agricultural
Sciences, Sweden
Muhammad Jamil,
King Abdullah University of Science
and Technology, Saudi Arabia

*Correspondence:

Kahi Ngugi
kahi@uonbi.ac.ke
Damaris A. Odeny
d.odeny@cgiar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 25 February 2021

Accepted: 03 June 2021

Published: 08 July 2021

Citation:

Muchira N, Ngugi K, Wamalwa LN, Avosa M, Chepkorir W, Manyasa E, Nyamongo D and Odeny DA (2021) Genotypic Variation in Cultivated and Wild Sorghum Genotypes in Response to *Striga hermonthica* Infestation. *Front. Plant Sci.* 12:671984. doi: 10.3389/fpls.2021.671984

Striga hermonthica is the most important parasitic weed in sub-Saharan Africa and remains one of the most devastating biotic factors affecting sorghum production in the western regions of Kenya. Farmers have traditionally managed *Striga* using cultural methods, but the most effective and practical solution to poor smallholder farmers is to develop *Striga*-resistant varieties. This study was undertaken with the aim of identifying new sources of resistance to *Striga* in comparison with the conventional sources as standard checks. We evaluated 64 sorghum genotypes consisting of wild relatives, landraces, improved varieties, and fourth filial generation (F₄) progenies in both a field trial and a pot trial. Data were collected for days to 50% flowering (DTF), dry panicle weight (DPW, g), plant height (PH, cm), yield (YLD, t ha⁻¹), 100-grain weight (HGW, g), overall disease score (ODS), overall pest score (OPS), area under *Striga* number progress curve (ASNPC), maximum above-ground *Striga* (NS_{max}), and number of *Striga*-forming capsules (NSFC) at relevant stages. Genetic diversity and hybridity confirmation was determined using Diversity Arrays Technology sequencing (DART-seq). Residual heterosis for HGW and NS_{max} was calculated as the percent increase or decrease in performance of F₄ crossover midparent (MP). The top 10 best yielding genotypes were predominantly F₄ crosses in both experiments, all of which yielded better than resistant checks, except FRAMIDA in the field trial and HAKIKA in the pot trial. Five F₄ progenies (ICSVIII IN × E36-1, LANDIWHITE × B35, B35 × E36-1, F6YQ212 × B35, and ICSVIII IN × LODOKA) recorded some of the highest HGW in both trials revealing their stability in good performance. Three genotypes (F6YQ212, GBK045827, and F6YQ212xB35) and one check (SRN39) were among the most resistant to *Striga* in both trials. SNPs generated from DART-seq grouped the genotypes into three major clusters, with all resistant checks grouping in the same cluster except N13. We identified more resistant and high-yielding genotypes than the conventional checks, especially among the F₄ crosses, which should be promoted for adoption by farmers. Future studies will need to look for more diverse sources of *Striga* resistance and pyramid different mechanisms of resistance into farmer-preferred varieties to enhance the durability of *Striga* resistance in the fields of farmers.

Keywords: witchweed, residual heterosis, DART-seq, CWR, pre-breeding, *Striga hermonthica*

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is a diploid ($2n = 2x = 20$) cereal grass of the Gramineae family native to Africa (Doggett, 1988). It is the fifth most important cereal globally (Kiprotich et al., 2015) and a major staple food for more than 300 million inhabitants of Africa (Kidanemariam et al., 2018). In Kenya, sorghum is ranked second after maize (*Zea mays* L.) in tonnage and production area, which is approximately 117,000 ha (FAOSTAT, 2016). Drought stress and poor soil fertility are the major abiotic factors affecting sorghum production in semi-arid areas (Ejeta and Knoll, 2007). Biotic stresses include diseases such as anthracnose (*Colletotrichum graminicola*) (Marley et al., 2005), leaf blight (*Exserohilum turcicum*) (Beshir et al., 2015), and the parasitic weed (*Striga hermonthica*).

The genus *Striga* comprises over 30 species, of which *S. hermonthica*, also known as the purple witchweed, is the most important in sub-Saharan Africa. *S. hermonthica* parasitizes several major cereal crops including maize, sorghum, rice (*Oryza sativa* L.), finger millet [*Eleusine coracana* (L.) Gaertn.], and pearl millet [*Pennisetum glaucum* (L.) R. Br.]. It remains one of the most devastating biotic factors affecting sorghum production in the western regions of Kenya (Khan et al., 2006) often characterized by low fertility and high moisture stress. The weed germinates on stimulation by a strigolactone (Bouwmeester et al., 2019; Aliche et al., 2020) induced by the host, or in some cases, non-host plants. The germinated *Striga* then attaches to the roots of the host plants, using a special invasive organ, the haustorium (Yoshida and Shirasu, 2009). The haustorium enables uptake of water and nutrients from the host plants for the growth and development of *Striga*, as well as the introduction of phytotoxins to the host (Van Hast et al., 2000). Consequently, the growth and development of the host plants become severely affected resulting in yield losses of up to 100% (Kim et al., 2002; Ejeta, 2007).

An adult *Striga* plant can produce up to 100,000 tiny seeds that can survive in the soil for 20 years or more (Pieterse and Pesch, 1983; Gurney et al., 2005), making it extremely difficult to control. Previous studies have reported *Striga* seed and plant densities in western Kenya at about 1,188 seeds per mature *Striga* seed capsule (Van Delft et al., 1997) and about 14 plants per m² (MacOpiyo et al., 2010), respectively. In Kenya, the three crops most devastated by *S. hermonthica* are maize, finger millet, and sorghum. Traditionally, farmers have managed *Striga* in sorghum fields using cultural and mechanical methods including hand weeding (Frost, 1994), intercropping (Aasha et al., 2017), and crop rotations (Oswald and Ransom, 2001) with edible legumes such as common bean (*Phaseolus vulgaris* L.), pigeon pea [*Cajanus cajan* (L.) Millsp.], and mung bean [*Vigna radiata* (L.) R. Wilczek].

Pathogenic isolates of *Fusarium oxysporum* f. sp. *Strigae* have been reported to be effective as bioherbicides, especially when integrated with other control practices (Rebeka et al., 2013). The push-pull technology that involves the intercropping of cereals with a trap crop (pull), usually Napier grass (*Pennisetum purpureum*), and a forage legume, usually desmodium (*Desmodium* spp.), as a push crop (Khan et al., 2011) has resulted in low adoption due to the lack of use for desmodium

by farmers. “Suicidal death” of *Striga*, which is achieved by inducing the germination of *Striga* by non-host legumes, has been employed in the reduction of *Striga* seed banks (Rubiales and Fernández-Aparicio, 2012). Chemical control has been tested in maize (Menkir et al., 2010) and sorghum (Bouréma et al., 2005; Dembélé et al., 2005; Tuinstra et al., 2009), but it is not environmentally friendly besides being unaffordable for the average sorghum farmer in Kenya. The most effective and practical solution for the smallholder sorghum farmers is to develop *Striga*-resistant sorghum varieties.

Sorghum germplasm screening against *Striga* is the first step toward the identification of *Striga*-resistant genotypes. Resistance has been reported among cultivated sorghum varieties including N13 (Haussmann et al., 2004), SRN39, FRAMIDA, and IS9830 (Ezeaku, 2005; Rodenburg et al., 2005). The resistance mechanism in N13 is a hypersensitive reaction characterized by thickening of the cell wall and silica deposition that limits xylem–xylem connection with the host plants (Maiti et al., 1984). SRN39, on the other hand, is known to harbor pre-attachment resistance that results in the production of a low germination stimulant, orobanchol (Satish et al., 2012; Mohamed et al., 2016). N13 and SRN39 have been used extensively as sources of resistance (Hess and Ejeta, 1992; Ngugi et al., 2015; Yohannes et al., 2015; Ali et al., 2016), and the quantitative trait loci (QTLs) responsible for resistance have been mapped (Haussmann et al., 2004; Satish et al., 2012). The outcrossing nature of *Striga* that results in different ecotypes with mixed responses to different genotypes (Fantaye, 2018) would require the pyramiding of multiple alleles from diverse sources into farmer-preferred varieties if the resistance were to be durable. Crop wild relatives (CWRs) and landraces of sorghum have been reported with significantly higher resistance to *Striga* than N13 and SRN39 (Mbuvi et al., 2017; Mallu et al., 2021). Such reports provide strong justification for more screening of CWRs and landraces toward the identification of additional sources of resistance to *Striga*.

Recommended methodologies for effective field screening include inoculation with *Striga* seeds, appropriate experimental designs with sufficient replications, quantitative data scoring, and inclusion of susceptible and resistant checks at regular intervals (Haussmann et al., 2000b; Rodenburg et al., 2005). A quantitative measure such as “area under *Striga* number progress curve” (ASNPC) alongside *Striga* count, *Striga* vigor, and grain yield has been used in past studies (Haussmann et al., 2012; Abate et al., 2017) with great success. The aim of this study was to screen for novel sources of resistance to *Striga* using both wild and landrace sorghum accessions, improved sorghum varieties, selected F₄ progenies, and known *Striga*-resistant sources, such as N13, FRAMIDA, HAKIKA, IS9830, and SRN39, as checks.

MATERIALS AND METHODS

Plant Material and Experimental Design

Sixty-four sorghum genotypes (Table 1) consisting of 17 wild relatives, 8 landraces, 13 improved varieties (high yielding), and 26 F₄ progenies of selected parents were used in this

TABLE 1 | Sorghum genotypes used, their sources, and classification.

Genotype	Source	Classification	Species
1. GBK 044058	GeRRI	Wild	<i>Sorghum</i> sp.
2. GBK 044336	GeRRI	Wild	<i>Sorghum</i> sp.
3. GBK 048922	GeRRI	Wild	<i>Sorghum</i> sp.
4. GBK 047293	GeRRI	Wild	<i>Sorghum arundinaceum</i> (Desv.) Stapf
5. GBK 048916	GeRRI	Wild	<i>Sorghum</i> sp.
6. GBK 016085	GeRRI	Wild	<i>Sorghum arundinaceum</i> (Desv.) Stapf
7. GBK 048917	GeRRI	Wild	<i>Sorghum</i> sp.
8. GBK 016114	GeRRI	Wild	<i>Sorghum sudanense</i> (Piper) Stapf
9. GBK 044063	GeRRI	Wild	<i>Sorghum</i> sp.
10. GBK 048156	GeRRI	Wild	<i>Sorghum arundinaceum</i> (Desv.) Stapf
11. GBK 016109	GeRRI	Wild	<i>Sorghum arundinaceum</i> (Desv.) Stapf
12. GBK 044120	GeRRI	Wild	<i>Sorghum</i> sp.
13. GBK 040577	GeRRI	Wild	<i>Sorghum arundinaceum</i> (Desv.) Stapf
14. GBK 048921	GeRRI	Wild	<i>Sorghum</i> sp.
15. GBK 044448	GeRRI	Wild	<i>Sorghum</i> sp.
16. GBK 045827	GeRRI	Wild	<i>Sorghum purpureosericeum</i> (Hochst. ex A. Rich.) Asch. and Schweinf.
17. GBK 048152	GeRRI	Wild	<i>Sorghum arundinaceum</i> (Desv.) Stapf
18. GBK 044065	GeRRI	Landrace	<i>Sorghum</i> sp.
19. GBK 043565	GeRRI	Landrace	<i>Sorghum arundinaceum</i> (Desv.) Stapf
20. GBK 044054	GeRRI	Landrace	<i>Sorghum alnum</i> Parodi
21. OKABIR	ICRISAT	Landrace	<i>Sorghum bicolor</i>
22. IS9830*	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
23. HAKIKA*	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
24. AKUOR-ACHOT	ICRISAT	Landrace	<i>Sorghum bicolor</i>
25. LODOKA†	ICRISAT	Landrace	<i>Sorghum bicolor</i>
26. E36-1†	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
27. B35†	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
28. N13*	ICRISAT	Landrace	<i>Sorghum bicolor</i>
29. SRN39*	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
30. KARIMTAMA-1	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
31. GADAM	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
32. F6YQ212	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
33. MACIA	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
34. FRAMIDA*	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
35. KAT/ELM/2016 PL82 KM32-2	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
36. KAT/ELM/2016 PL1 SD15	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
38. ICSVIII_IN	ICRISAT	Improved variety	<i>Sorghum bicolor</i>
39. OKABIR × AKUOR-ACHOT	UON	F4 Population	<i>Sorghum bicolor</i>
40. AKUOR-ACHOT × ICSVIII_IN	UON	F4 Population	<i>Sorghum bicolor</i>
41. B35 × AKUOR-ACHOT	UON	F4 Population	<i>Sorghum bicolor</i>
42. B35 × E36-1	UON	F4 Population	<i>Sorghum bicolor</i>
43. B35 × F6YQ212	UON	F4 Population	<i>Sorghum bicolor</i>
44. B35 × ICSVIII_IN	UON	F4 Population	<i>Sorghum bicolor</i>
45. B35 × LANDIWHITE	UON	F4 Population	<i>Sorghum bicolor</i>
46. B35 × LODOKA	UON	F4 Population	<i>Sorghum bicolor</i>
47. E36-1 × MACIA	UON	F4 Population	<i>Sorghum bicolor</i>
48. F6YQ212 × B35	UON	F4 Population	<i>Sorghum bicolor</i>
49. F6YQ212 × LODOKA	UON	F4 Population	<i>Sorghum bicolor</i>
50. IBUSAR × E36-1	UON	F4 Population	<i>Sorghum bicolor</i>
51. IBUSAR × LANDIWHITE	UON	F4 Population	<i>Sorghum bicolor</i>
52. IBUSAR × ICSVIII_IN	UON	F4 Population	<i>Sorghum bicolor</i>

(Continued)

TABLE 1 | Continued

Genotype	Source	Classification	Species
53. ICSVIII_IN × B35	UON	F4 Population	<i>Sorghum bicolor</i>
54. ICSVIII_IN × E36- 1	UON	F4 Population	<i>Sorghum bicolor</i>
55. ICSVIII_IN × LANDIWHITE	UON	F4 Population	<i>Sorghum bicolor</i>
56. ICSVIII_IN × LODOKA	UON	F4 Population	<i>Sorghum bicolor</i>
57. ICSVIII_IN × MACIA	UON	F4 Population	<i>Sorghum bicolor</i>
58. LANDIWHITE × B35	UON	F4 Population	<i>Sorghum bicolor</i>
59. LANDIWHITE × MACIA	UON	F4 Population	<i>Sorghum bicolor</i>
60. LODOKA × ICSVIII_IN	UON	F4 Population	<i>Sorghum bicolor</i>
61. LODOKA × LANDIWHITE	UON	F4 Population	<i>Sorghum bicolor</i>
62. LODOKA × OKABIR	UON	F4 Population	<i>Sorghum bicolor</i>
63. OKABIR × B35	UON	F4 Population	<i>Sorghum bicolor</i>
64. OKABIR × ICSVIII_IN	UON	F4 Population	<i>Sorghum bicolor</i>

GeRRI, Genetic Resources Research Institute-Kenya; F₄, fourth filial generation; ICRISAT, International Crops Research Institute for the Semi-Arid Tropics; UON, University of Nairobi.

**Striga*-resistant checks.

† Staygreen source.

study. The parental lines of the crosses included five improved varieties (B35, E36-1, MACIA, ICSVIII IN, and F6YQ212) and five landraces (LODOKA, OKABIR, IBUSAR, AKUOR-ACHOT, and LANDIWHITE). Successful crosses were selected morphologically at F₁ and advanced to F₄ using the bulk-population method.

The field and pot trials were set up during the long rainy season of 2019 at a *S. hermonthica* hotspot in Alupe, Busia County, Western Kenya. Alupe is located on 34° 07' 28.6" E and 00° 30' 10.1" N with an annual rainfall range of 1,100–1,450 mm and daily mean temperatures of 24°C. The area has a bimodal rainfall pattern of long and short rains. Both experiments were laid out in a square lattice design with three replications, each block consisting of eight plots. The field experiment was planted in a *Striga*-infested field with a spacing of 75 cm between rows and 20 cm between plants in the row. Each row contained 21 plants. *Striga* inoculum was prepared by mixing 5 kg of sand with 15 g of *Striga* seeds that had been harvested from the same location in the previous season. A supplemental *Striga* inoculum of 15 g was spread along each row during planting to improve the consistency of *Striga* seed load across the plot. Phosphorus (P) was applied at the rate of 90 Kg ha⁻¹ after thinning, while nitrogen (N) was applied at the rate of 92 kg ha⁻¹ when the plants were 45–50 cm tall, which was around 30 days after germination. Insect pests, especially fall armyworm (*Spodoptera frugiperda*) and cutworms (*Agrotis* spp., *Spodoptera* spp., and *Schizonycha* spp.), were controlled using Voliam Targo® SC (Syngenta Crop Protection AG, Switzerland) containing active ingredients such as chlorantraniliprole and abamectin. The field experiment was purely rainfed.

For the pot experiment, pots of 30 cm diameter were filled with 20 kg of *Striga*-free soil obtained from *Striga*-free field. Each pot contained four plants and was used to represent a plot. The pot experiment was set up in the field alongside the field experiment and was not under any shelter. The pot experiment was rainfed as much as possible but due to the restricted pot size

and high levels of evaporation from the pots, watering was done only when the plants were close to the permanent wilting point. *Striga* inoculation was done by adding 5 g of *Striga* inoculum to each pot. The application of fertilizer and insect pest control was carried out as already described.

Agronomic and *Striga* Data Collection

The data on agronomic traits were collected from six randomly selected plants from each plot, while the data on *Striga* response traits were collected per plot. The agronomic data were collected for days to 50% flowering (DTF), dry panicle weight (DPW, g), plant height (PH, cm), yield (YLD, t ha⁻¹), and 100-grain weight (HGW, g), at relevant stages as per the recommendations of the International Board for Plant Genetic Resources (IBPGR) and the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) (1993). The date of first *Striga* emergence was recorded followed by *Striga* count at 2-week intervals until maturity. The number of *Striga* plants forming capsules (NSFC) per plot was recorded at 105 days after sowing. We used the overall disease score (ODS) and the overall pest score (OPS) on a scale of 1–9 to account for any diseases and pests observed, ranging from leaf blight (*Helminthosporium turcicum*), ladder leaf spot (*Cercospora fuscimaculans*), zonate leaf spot (*Gloeocercospora sorghi*), anthracnose (*C. graminicola*), spider mites (*Oligonychus pratensis* and *Tetranychus urticae*), sorghum midge (*Contarinia sorghicola*), and sorghum shoot fly (*Atherigona soccata*).

ANOVA and *Striga* Data Analysis

The maximum above-ground *Striga* (NS_{max}) was calculated as suggested by Rodenburg et al. (2006). The ASNPC was calculated according to the formula suggested by Haussman et al. (2000) as follows:

$$ASNPC = \sum_{i=0}^{n-1} \left[\frac{Y_i + Y_{(i+1)}}{2} \right] (t_{(i+1)} - t_i)$$

where n is the number of *Striga* assessment dates, Y_i is the *Striga* count at the i th assessment date, and t_i is the number of days after sowing at the i th assessment date.

ANOVA and means for quantitative traits were generated in the alpha lattice design using Genstat software version 19.1 (VSN International, 2017). Treatment means were compared using Fisher's protected least significant differences at $P \leq 0.05$. The estimates of phenotypic and genotypic variance and the genotypic and phenotypic coefficients of variation were performed based on the formula proposed by Syukur et al. (2012).

Genotypic variance:

$$\sigma_g^2 = \frac{MS_g - MS_e}{r}$$

Phenotypic variance:

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2$$

where σ_g^2 = genotypic variance, σ_p^2 = phenotypic variance, σ_e^2 = environmental variance (i.e., error mean square from the ANOVA), MS_g = mean square of genotypes, MS_e = error mean square, and r = number of replications.

Genotypic coefficient of variation:

$$[GCV] = \left[\frac{\sqrt{\sigma_g^2}}{\bar{x}} \right] \times 100$$

Phenotypic coefficient of variation:

$$[PCV] = \left[\frac{\sqrt{\sigma_p^2}}{\bar{x}} \right] \times 100$$

where σ_g^2 = genotypic variance, σ_p^2 = phenotypic variance, and \bar{x} is grand mean of a character.

Simple linear correlation coefficients were calculated to understand the relationship among the studied agronomic traits for each trial according to the formula given below:

$$P_{X,Y} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

where cov is the covariance, σ_x is the SD of x , and σ_y is the SD of Y .

Phenotypic correlations across the field and pot trials were estimated, and correlation plots were drawn by using R Version 4.0.4 according to the formula described by (Hallauer et al., 2010):

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

where $r_{x,y}$ is the correlation coefficient of each trait between the two sites, x (field trial), and y (pot trial); \bar{x} and \bar{y} are the means of the values of each of the traits in the field and

pot trials, respectively. The significance of linear relationships in phenotypic correlation coefficients across the two trials was compared with r -coefficient values and the associated degrees of freedom ($n = 2$), at the probability levels of $P = 0.05$, 0.01, and 0.001.

Heritability Estimates

Estimations of broad-sense heritability (H^2) for all traits were calculated based on parental and family means, respectively, according to the formula described by Allard (1960):

$$H^2 bs = \left[\frac{\sigma_g^2}{\sigma_p^2} \right] \times 100$$

$H^2 bs$ = heritability in broad sense, σ_g^2 = genotypic variance, and σ_p^2 = phenotypic variance. H^2 scores were classified according to Robinson et al. (1949) as follows: 0–30% = low, 30–60% = moderate, and >60% = high.

Genotyping, Genetic Relatedness, and Confirmation of True Crosses

Molecular data of all the 37 parental genotypes in Table 1, which consisted of 17 wild accessions, 8 landraces, and 12 improved varieties, were generated for the genetic diversity analysis. DNA extraction, genotyping, and filtering of raw SNPs were performed as described by Ochieng et al. (2020). A neighbor-joining (NJ) dendrogram was generated using the Trait Analysis by aSSociation, Evolution, and Linkage (TASSEL) software version 5.2.67 (Bradbury et al., 2007). To undertake the hybridity confirmation, DNA was extracted from additional 115 F_4 progenies that were representative of all the crosses, bringing the total number of individuals genotyped to 153. The filtered SNP variant call file for each of the genotyped F_4 progenies was parsed through the GenosToABHPlugin in TASSEL 5.2.67 alongside the corresponding parents to obtain informative biallelic SNPs in the ABH format (i.e., female parent alleles as “A,” male parent allele as “B,” and heterozygotes as “H”).

Comparing the Agronomic Performance of Parents and Progenies Under *Striga* Field Trial Conditions

Crosses involving parents that failed to germinate were excluded from the analysis. The percentage increase or decrease in the performance of F_4 crossover midparent (MP) was calculated to observe the residual heterotic effects for HGW and NS_{max} . The average F_4 -values per cross were used for the estimation of residual heterosis expressed in percentage over MP as described by Turner (1953).

where,

$$MP\text{-value} = (P1 + P2)/2$$

$$\text{Residual heterosis} = [(F_4 - MP)/MP] \times 100$$

TABLE 2 | Mean squares of agronomic traits measured under field trial and pot trial.

	SOV	HGW	YLD	DTF	PH
Combined	Environment	21.65***	2,779.93***	4,555.74***	545.20ns
Field trial	Reps	2.21*	1.05ns	81.84ns	2,355.90***
	Genotype	2.34***	4.40***	319.67***	6,421.40***
	Parents (Par)	2.04***	2.69***	213.02***	5,566.20***
	Progenies (Pro)	2.31***	5.41***	465.04***	6,655.70***
	Par × Pro	13.16***	42.86***	98.2ns	27,930***
	Residual	0.75	0.76	49.87	366.62
Pot trial	Reps	2.1421ns	32.81ns	21.87ns	205.80ns
	Genotype	1.99***	164.89***	177.24***	6,192.20***
	Parents (Par)	3.92***	118.66***	140.09***	6,860***
	Progenies (Pro)	1.67ns	154.39***	228.61***	5,356.80***
	Par × Pro	17.84***	2,081.54***	335.70***	592ns
	Residual	1.30	40.69	45.17	575.75

HGW, 100-grain weight; YLD, yield; DTF, days to 50% flowering; PH, plant height.

*Significant at $P < 0.05$.

***Significant at $P < 0.001$.

ns, non-significant.

TABLE 3 | Mean squares of *Striga*-, disease-, and pest-related traits measured under field trial and pot trial.

	SOV	ASNPC	NS _{max}	NSFC	ODS	OPS
Combined	Environment	35,611,065***	71,590.40***	10,788.95***	1.80**	4.9515ns
Field trial	Reps	2,289,177ns	3,319*	384.8*	0.43ns	2.59***
	Genotype	7,332,224***	4,620***	570***	3.68***	3.96***
	Parents (Par)	7,863,173***	5,402***	742.2***	3.56***	2.89***
	Progenies (Pro)	6,203,111***	3,676***	340.4***	3.69***	5.30***
	Par × Pro	292,494***	880ns	318.7ns	7.41***	4.44***
	Residual	1,860,949	1,333.39	135.32	0.58	0.54
Pot trial	Reps	491,477***	349.13***	2.59ns	0.35ns	4.10**
	Genotype	333,365***	73.39***	8.64***	5.28***	2.59***
	Parents (Par)	448,828***	83.86***	10.02***	4.53***	2.57***
	Progenies (Pro)	169,309***	57.10***	6.73**	4.36***	2.27***
	Par × Pro	162,659*	93.81***	5.54ns	58.79***	10.98***
	Residual	88,840	11.18	3.96	1.12	1.04

ASNPC, area under *Striga* number progress curve; NS_{max}, maximum above-ground *Striga*; NSFC, number of *Striga*-forming capsules; ODS, overall disease score; OPS, overall pest score.

*Significant at $P < 0.05$.

**Significant at $P < 0.01$.

***Significant at $P < 0.001$.

ns, non-significant.

RESULTS

Field Trial and Pot-Screening Trial

All the trait means from the field and pot trials are provided in **Supplementary Table 1**. The two environments had an effect on all traits except PH (**Table 2**) and OPS (**Table 3**). Significant differences ($P \leq 0.001$) were observed in the agronomic performance of genotypes for all traits across the field and pot trials (**Table 2**). There were also significant differences ($P \leq 0.001$) in the performance of the parental lines when compared with the progenies, except for DTF in the field trial and PH

in the pot trial (**Table 2**). We observed more consistency across replications in the pot trial for all agronomic traits than in the field trial, where differences across replications were observed for HGW and PH (**Table 2**).

Significant differences ($P < 0.001$) were observed between genotypes for all *Striga*-related traits, as well as for ODS and OPS (**Table 3**). The performance of parents against their progenies also revealed significant differences ($P < 0.001$) for ASNPC, ODS, and OPS. More consistency across replications was observed in the field trial for ASNPC and ODS and in the pot trial for NSFC and ODS (**Table 3**). Yield-related traits (i.e., YLD and

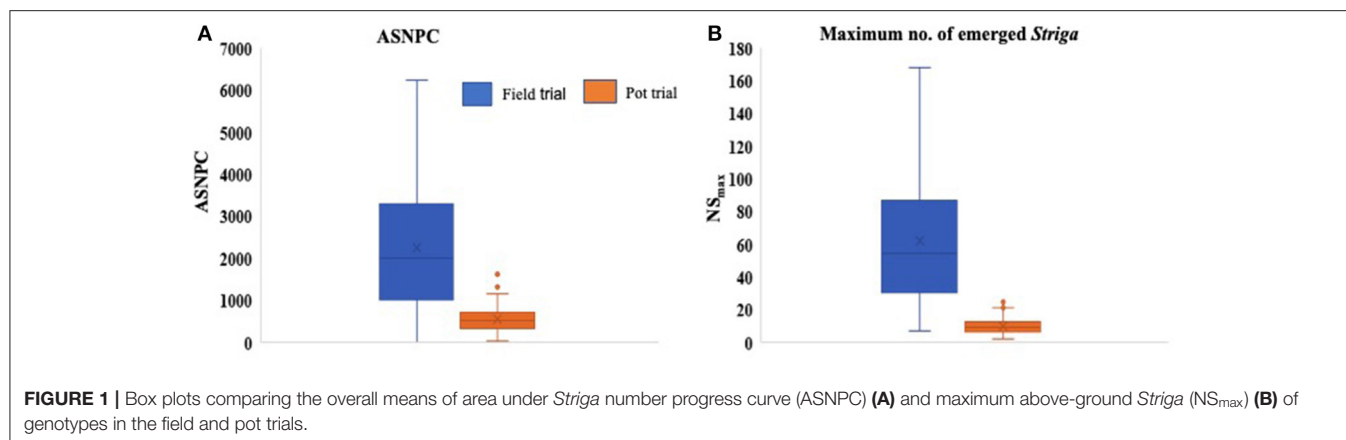


FIGURE 1 | Box plots comparing the overall means of area under *Striga* number progress curve (ASNPC) (A) and maximum above-ground *Striga* (NS_{max}) (B) of genotypes in the field and pot trials.

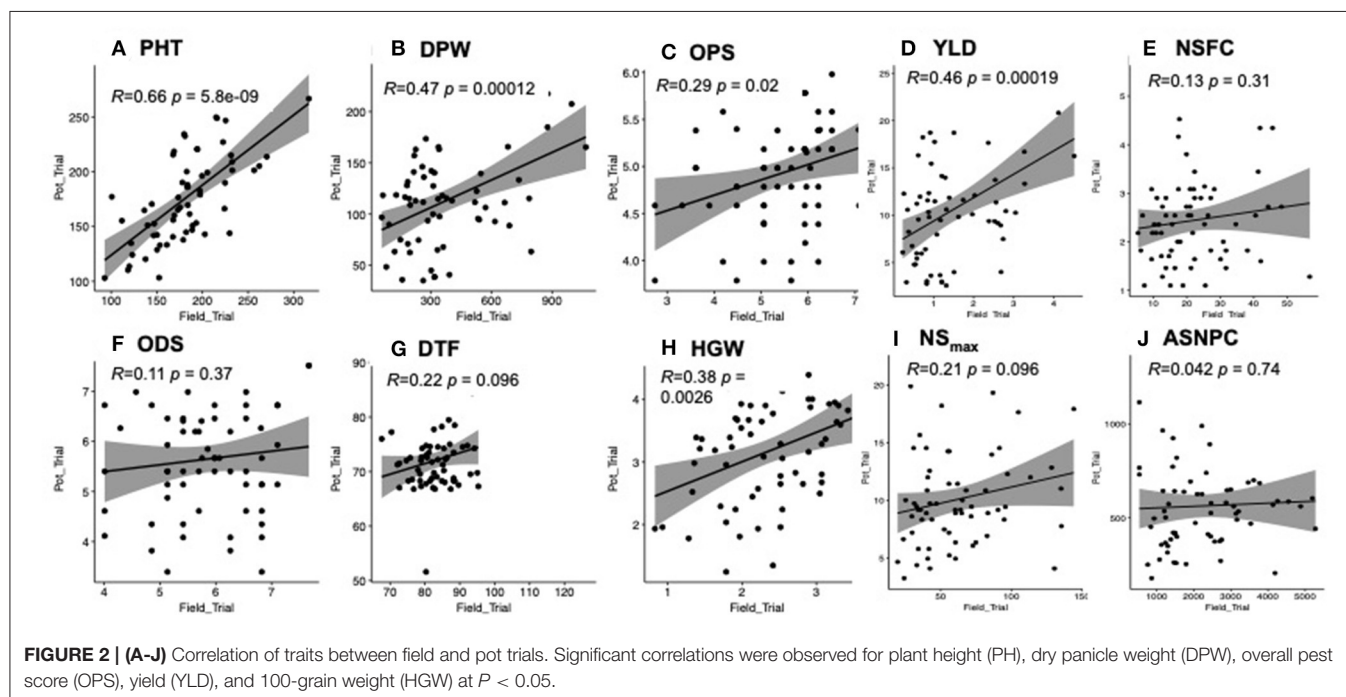


FIGURE 2 | (A–J) Correlation of traits between field and pot trials. Significant correlations were observed for plant height (PH), dry panicle weight (DPW), overall pest score (OPS), yield (YLD), and 100-grain weight (HGW) at $P < 0.05$.

HGW) were consistently higher in the pot trial than in the field trial, while *Striga*-related traits (i.e., ASNPC and NS_{max}) were lower in the pot trial in comparison with the field trial (Figure 1).

Trait Correlations and Heritability

We observed positive and significant correlations between the yield-related data collected from the field and pot trials for all traits (i.e., PH, DPW, YLD, and HGW) except DTF (Figure 2). There was a weak non-significant correlation between the two trials for all the biotic stress-related traits (i.e., NSFC, ASNPC, NS_{max}, and ODS) except OPS (Figure 2).

We also looked at trait correlations within each trial and recorded more significant trait correlations ($P < 0.05$) in the pot trial (21) than in the field trial (13) (Supplementary Table 1). Yield-related traits (i.e., HGW, YLD, DTF, and PH) were

negatively correlated with ASNPC, NS_{max}, NSFC, ODS, and OPS, in both field and pot trials, although the correlation was largely non-significant (Supplementary Table 1). The highest positive significant correlations in the field trial were recorded between ASNPC and NS_{max} ($r = 0.83$; $P < 0.001$), ASNPC and NSFC ($r = 0.77$; $P < 0.001$), and NS_{max} and NSFC ($r = 0.80$; $P < 0.001$). The same traits were also highly positively correlated in the pot trial with comparable correlation values of $r = 0.84$ (i.e., ASNPC and NS_{max}) and $r = 0.73$ (i.e., NSFC and ASNPC), except for the correlation between NSFC and NS_{max} ($r = 0.55$).

Parents recorded high heritability values for all traits except DTF under the pot trial (Figure 3, Supplementary Table 1). Progenies displayed relatively lower heritability values for most traits in comparison with the parents, except for DTF in both environments (Figure 3).

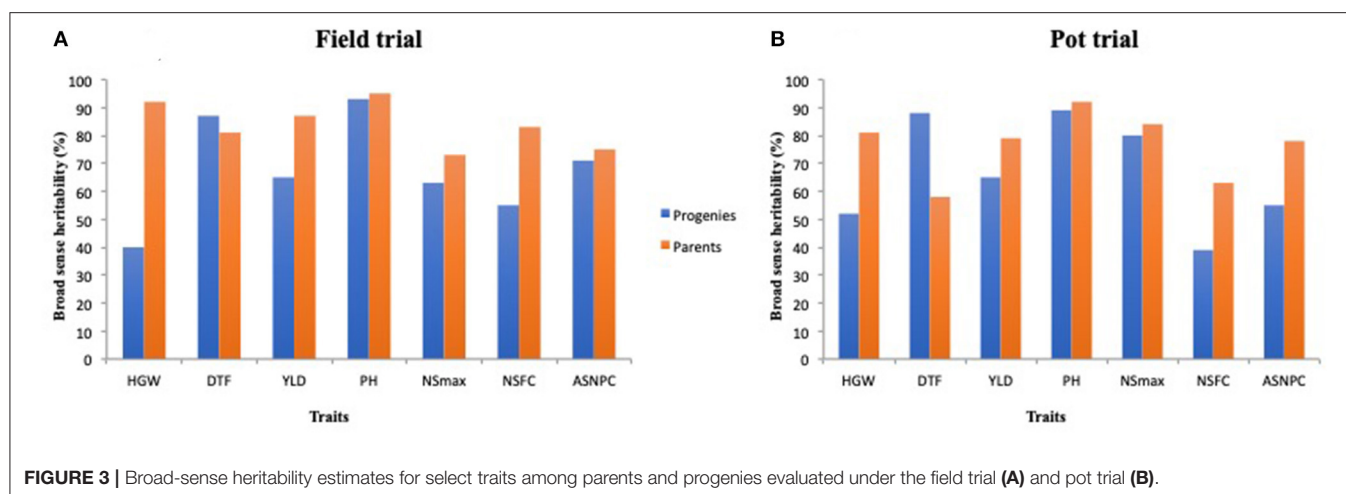


FIGURE 3 | Broad-sense heritability estimates for select traits among parents and progenies evaluated under the field trial (A) and pot trial (B).

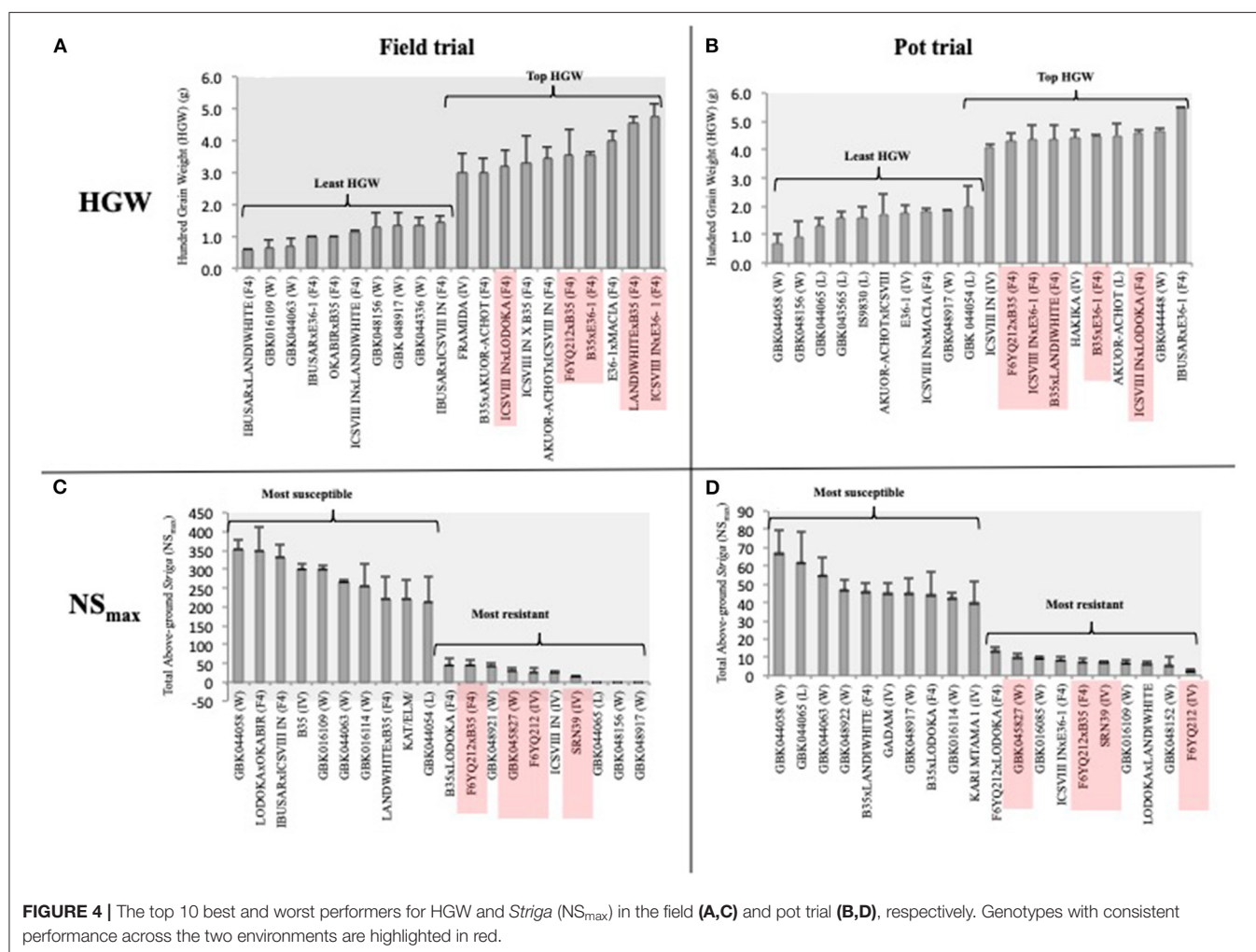


FIGURE 4 | The top 10 best and worst performers for HGW and *Striga* (NS_{max}) in the field (A,C) and pot trial (B,D), respectively. Genotypes with consistent performance across the two environments are highlighted in red.

Agronomic Performance of the Germplasm Under Field and Pot Trials

We used HGW rather than YLD to compare the yield performance of genotypes between the two trials, given the

differences in the trial conditions that would bias the total yield comparisons. Of note, 9 and 6 out of the 10 genotypes with the highest HGW in the field and pot trials, respectively, were F₄ progenies (Figures 4A,B). Five (ICSVIII IN × E36-1,

LANDIWHITE \times B35, B35 \times E36-1, F6YQ212 \times B35, and ICSVIII IN \times LODOKA) of the F₄ progenies with the highest HGW were common in both trials (**Figures 4A,B**), revealing their potential stability for the trait. FRAMIDA and HAKIKA, which are both *Striga*-resistant and improved varieties, were the only resistant checks among the top 10 genotypes recording the highest HGW in the field and pot trials, respectively (**Figures 4A,B**). A landrace, AKUOR-ACHOT, and a wild accession, GBK044448, were also among the top 10 genotypes with high HGW of 4.5 and 4.7 g, respectively, in the pot trial (**Figure 4B**) but not in the field trial (**Supplementary Table 1**). Most of the genotypes with the lowest HGW were wild accessions or landraces in both trials, although some improved varieties (i.e., E36-1 in pot trial) and F₄ progenies also fell into this category (**Figures 4A,B**).

Response of Genotypes to *Striga* Under Field and Pot Trials

Despite the poor correlation of *Striga*-related traits between pot and field trials, we observed stability in the performance of four (F6YQ212, SRN39, F6YQ212 \times B35, and GBK045827) out of the 10 top *Striga*-resistant genotypes in both field and pot trials (**Figures 4C,D**) according to their NS_{max} ranking (**Figures 4C,D**). SRN39, a stable source of resistance to *Striga*, was the only resistant check among the topmost resistant genotypes in both trials (**Figures 4C,D**). Although the F₄ progenies were the dominant best performers for HGW, they were the minority genotypes among the top 10 most *Striga*-resistant genotypes in both field and pot trials. F6YQ212, which showed resistance to *Striga* in both trials, was the common parent in two out of the three F₄ progenies with resistance to *Striga*, suggesting it would be a good source of stable *Striga* resistance for future crosses. B35, a drought-tolerant variety, which was among the top 10 most *Striga*-susceptible genotypes in the field trial, was the common parent in two out of the three top *Striga*-resistant progenies in both trials (**Figures 4C,D**). The only genotype that was recorded among the top performers in both trials for HGW and *Striga* resistance (NS_{max}) was a cross between the *Striga*-resistant variety, F6YQ212, and the drought-tolerant variety, B35 (**Figure 4**).

Three genotypes that included two wild (GBK048156 and GBK048917) and a landrace (GBK044065) recorded no *Striga* germination in the field trial but supported the germination of significant amounts of *Striga* in the pot trial (**Supplementary Table 1**). A wild accession, GBK044058, was the most susceptible genotype to *Striga* in both trials. All genotypes tested for the *Striga* germination in the pot trial supported the germination of at least three *Striga* plants in at least one replicate (**Supplementary Table 1**). However, genotype GBK016109, which showed comparable resistance to *Striga* as SRN39 in the pot trial, was completely devastated in the field trial recording one of the worst performers (**Figures 4C,D**).

Genotype Relatedness and Confirmed Hybridity

A total of 26,291 raw SNPs were generated from DArT-seq of 64 genotypes (i.e., 17 wild, 8 landraces, 12 improved varieties,

and 27 F₄ progenies), of which 7,038 SNPs were retained after filtering. The NJ dendrogram resulted in three major clusters (**Figure 5**). The first cluster (A) (**Figure 5**) comprised of *Striga*-resistant genotypes including four resistant checks, namely, IS9830, SRN39, FRAMIDA, and HAKIKA. Other genotypes in cluster A were recorded as resistant to *Striga* in this study such as GBK048156 (field trial), GBK048152 (pot trial), F6YQ212 (field trial and pot trial), and GBK045827 (field trial and pot trial). The only susceptible genotype in this cluster was the improved variety KAT/ELM/2016PL1SD15. Cluster B comprised mostly of wild accessions and landraces. Two improved varieties, namely, B35 and MACIA, were also in cluster B, but different subclusters. The only *Striga*-resistant check in this cluster was N13, which was grouped in the same subcluster with the staygreen genotype, B35. Both B35 and N13 are known to have wild pedigrees.

Two wild genotypes (GBK016109 and GBK016085), which were among the top most *Striga*-resistant lines in the pot trial, and one landrace (GBK044065), which was among the most resistant lines in the field trial, were also grouped in cluster B (**Figure 5**). Cluster C was composed of four improved varieties, two landraces, and four wild accessions. Two genotypes (GBK048917 and ICSV III IN), which were also among the most resistant to *Striga* in the field trial, were also grouped under cluster C. Genotype E36-1, a well-known drought-tolerant (staygreen) material, was also grouped in cluster C alongside a landrace, GBK044054, which recorded one of the highest NS_{max} in the field trial (**Figure 4C**).

Crosses involving parental lines IBUSAR and LANDIWHITE were not included in the hybridity analysis as both parents failed to germinate in both trials. Hybridity of 16 F₄ progenies was confirmed using biallelic SNP markers ranging from 1,204 to 2,868 that had been called from DArT-seq (**Table 4**). The highest proportion (22–46%) of heterozygous alleles were recorded in the progenies of the cross B35 \times E36-1, while the lowest (<1%) were recorded in the crosses OKABIR \times AKUOR-ACHOT and LODOKA \times OKABIR. B35 and E36-1 are the improved varieties derived from wild backgrounds, while LODOKA and AKUOR-ACHOT are landraces.

Residual Heterosis for Yield and Resistance to *Striga* in the F₄ Progenies

A complete record of the MP and residual heterosis values in the field and pot trials for HGW and NS_{max} is shown in **Supplementary Table 2**. **Table 5** shows the residual heterosis values for HGW in each of the crosses in the field and pot trials ranked from the top to the lowest. Both the highest residual heterosis and the inbreeding depression for HGW were recorded in the field at 89.78% and –59.48%, respectively. LODOKA, a drought-tolerant landrace, was the common parent in the crosses with the highest residual heterosis in both trials (**Table 5**). Four (i.e., AKUOR-ACHOT \times ICSVIII IN, B35 \times AKUOR-ACHOT, ICSVIII IN \times MACIA, LODOKA \times ICSVIII IN) out of six crosses that recorded the inbreeding depression for HGW were consistent in both the field and pot trials (**Table 5**), suggesting they would be poor candidates for yield-related traits. Crosses involving ICSVIII IN, an improved variety, revealed some of the highest inbreeding depression for HGW in both field and

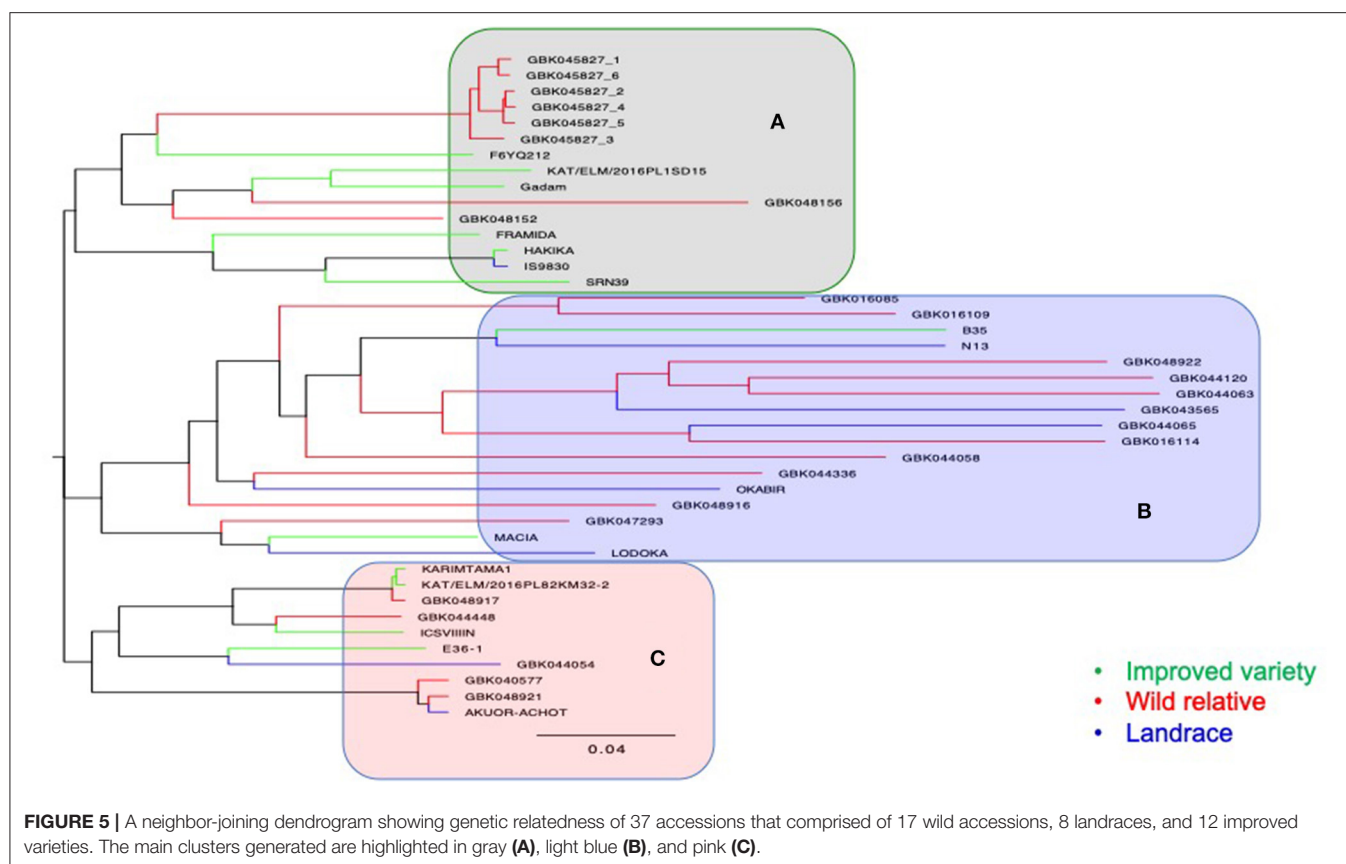


FIGURE 5 | A neighbor-joining dendrogram showing genetic relatedness of 37 accessions that comprised of 17 wild accessions, 8 landraces, and 12 improved varieties. The main clusters generated are highlighted in gray (A), light blue (B), and pink (C).

TABLE 4 | Confirmation of hybridity among the fourth filial generation (F_4) progenies using *Striga* number progress (SNP) markers.

Cross	Plants tested	Average no. of biallelic markers	Proportion of heterozygous alleles (%)
1. B35 × ICSVIII IN	6	2,806	<1–33
2. B35 × F6YQ212	6	2,774	15–43
3. B35 × AKUOR-ACHOT	6	2,699	1–4.5
4. E36-1 × MACIA	6	1,485	<1–29
5. F6YQ212 × LODOKA	5	1,947	2–13
6. B35 × ICSVIII IN	6	2,806	<1–33
7. B35 × E36-1	6	2,868	22–46
8. B35 × LODOKA	6	2,467	9–25
9. F6YQ212 × B35	5	2,695	<1–10
10. ICSVIII IN × E36-1	6	1,313	14–33
11. LODOKA × ICSVIII IN	6	1,825	<1–15
12. OKABIR × ICSVIII IN	6	2,339	<1–2
13. OKABIR × AKUOR-ACHOT	6	2,174	<1
14. LODOKA × OKABIR	6	2,357	<1
15. ICSVIII IN × MACIA	6	1,204	10–17
16. ICSVIII IN × LODOKA	4	1,764	4–13

pot trials (Table 5). In some cases, the highest residual heterosis recorded for HGW (86.19% for F6YQ212 × B35 and 52.94% for B35 × LODOKA) also corresponded to a low *Striga* count

of –65.13 and –79.20% (Tables 5, 6), making these crosses good candidates for the future development of *Striga*-tolerant, high-yielding varieties.

Table 6 shows the residual heterosis values for NS_{max} in each of the crosses in the field and pot trials. Crosses involving B35, a drought-tolerant improved variety, had some of the highest inbreeding depression values for NS_{max} , which is desirable as it resulted in more resistance to *Striga* (Table 6). Like in HGW, crosses involving ICSVIII IN performed worse than the MP for their resistance to *Striga*, indicating that this parent may not be a good choice for improving either yield or *Striga* resistance (Table 6).

DISCUSSION

Identification of Diverse High-Yielding *Striga*-Resistant Genotypes

The aim of this study was to identify new sources of *Striga* resistance in comparison with the conventional sources. The data collected found credible evidence that genotypes that were resistant to *Striga* had significantly lower ASNPC and NS_{max} -values in both trials. These parameters were considered alongside grain yield for the effective selection of superior *Striga*-resistant, as well as *Striga*-tolerant genotypes as recommended by Rodenburg et al. (2005). We identified three genotypes

TABLE 5 | Proportion of residual heterosis at F₄ for 100-grain weight (HGW) for the pot and field trials.

Field		Pot	
Crosses	HGW (%)	Crosses	HGW (%)
1. F6YQ212 × LODOKA	89.78	1. LODOKA × OKABIR	79.33
2. F6YQ212 × B35	86.19	2. E36-1 × MACIA	46.52
3. ICSVIII_IN × LODOKA	77.61	3. OKABIR × B35	31.03
4. B35 × LODOKA	52.94	4. B35 × E36-1	21.44
5. LODOKA × OKABIR	47.33	5. B35 × ICSVIII_IN	20.99
6. OKABIR × B35	41.44	6. ICSVIII_IN × E36-1	15.77
7. OKABIR × ICSVIII_IN	39.30	7. F6YQ212 × LODOKA	11.31
8. ICSVIII_IN × B35	30.25	8. OKABIR × ICSVIII_IN	10.14
9. B35 × F6YQ212	25.83	9. B35 × LODOKA	8.06
10. B35 × E36-1	24.53	10. OKABIR × AKUOR-ACHOT	4.52
11. E36-1 × MACIA	19.76	11. F6YQ212 × B35	3.33
12. ICSVIII_IN × E36-1	18.33	12. ICSVIII_IN × LODOKA	2.67
13. AKUOR-ACHOT × ICSVIII_IN	−0.50	13. B35 × F6YQ212	−8.33
14. B35 × AKUOR-ACHOT	−6.65	14. B35 × AKUOR-ACHOT	−21.58
15. B35 × ICSVIII_IN	−26.31	15. LODOKA × ICSVIII_IN	−22.32
16. ICSVIII_IN × MACIA	−27.05	16. ICSVIII_IN × B35	−37.46
17. LODOKA × ICSVIII_IN	−28.36	17. ICSVIII_IN × MACIA	−38.82
18. OKABIR × AKUOR-ACHOT	−59.48	18. AKUOR-ACHOT × ICSVIII_IN	−48.51

TABLE 6 | Proportion of residual heterosis at F₄ for maximum above-ground *Striga* (NS_{max}) for the pot and field trials.

Field		Pot	
Crosses	NS _{max} (%)*	Crosses	NS _{max} (%)*
1. B35 × LODOKA	−79.2	1. F6YQ212 × B35	−53.52
2. B35 × ICSVIII_IN	−71.21	2. ICSVIII_IN × B35	−45.44
3. B35 × AKUOR-ACHOT	−66.02	3. OKABIR × B35	−32.24
4. F6YQ212 × B35	−65.13	4. ICSVIII_IN × E36-1	−29.39
5. OKABIR × B35	−51.03	5. B35 × F6YQ212	−6.91
6. B35 × F6YQ212	−51.01	6. B35 × ICSVIII_IN	−5.4
7. ICSVIII_IN × B35	−49.36	7. LODOKA × OKABIR	−2.27
8. B35 × E36-1	−31.89	8. OKABIR × AKUOR-ACHOT	−1.81
9. OKABIR × AKUOR-ACHOT	−17.53	9. F6YQ212 × LODOKA	−0.12
10. ICSVIII_IN × LODOKA	15.48	10. B35 × AKUOR-ACHOT	1.57
11. ICSVIII_IN × E36-1	35.85	11. AKUOR-ACHOT × ICSVIII_IN	4.37
12. OKABIR × ICSVIII_IN	43.09	12. E36-1 × MACIA	21
13. F6YQ212 × LODOKA	62.09	13. ICSVIII_IN × MACIA	24.4
14. LODOKA × ICSVIII_IN	63.73	14. ICSVIII_IN × LODOKA	26.28
15. E36-1 × MACIA	75.88	15. B35 × E36-1	39.64
16. AKUOR-ACHOT × ICSVIII_IN	115.14	16. B35 × LODOKA	57.89
17. LODOKA × OKABIR	144.68	17. OKABIR × ICSVIII_IN	62.88
18. ICSVIII_IN × MACIA	289.77	18. LODOKA × ICSVIII_IN	136.78

*Values with negative percentage are the most desirable as they show a lower number of *Striga* plants than that of the midparent.

(F6YQ212, GBK045827, and F6YQ212 × B35) together with one check (SRN39) that were consistent in their response to *Striga* across both trials. SRN39 is known to harbor pre-attachment resistance that results in the production of a low germination stimulant, orobanchol (Satish et al., 2012; Mohamed et al., 2016).

However, SRN39 is not high yielding and, therefore, not preferred by farmers and has been used mainly as a donor of *Striga* resistance to other improved varieties. Of specific interest is the cross between improved varieties, F6YQ212 × B35, which recorded consistency in resistance to *Striga* and was also high

yielding. This particular cross is likely to perform well in the fields of farmers and would be a good genotype to advance for field trials.

The wild accession, GBK045827, which also showed consistency in *Striga* resistance across the two trials, did not only group together with F6YQ212 but also recorded comparable performance with F6YQ212 in both experiments. This observation strongly suggests that the resistance observed in F6YQ212 may have been introgressed from GBK045827. We know that *Striga* resistance is more abundant among wild relatives (Rich et al., 2004; Mbuvi et al., 2017), which tend to cross-pollinate with cultivated genotypes in open fields (Ohadi et al., 2017). Both F6YQ212 and GBK045827 clustered in group A with four resistant checks (FRAMIDA, HAKIKA, IS9830, and SRN39) but in a different subclade. This pattern of clustering suggests a narrow genetic base for *Striga*-resistant sources that are currently being used in breeding programs in Eastern Africa (Mohamed et al., 2010), except for N13. Genotype F6YQ212, therefore, provides a good alternative source of resistance as it was grouped in a different subclade. F6YQ212 has been previously screened for response to grain storage pests (Mwenda, 2019) but not for its resistance to *Striga*. Future studies will therefore need to determine the mode of resistance in F6YQ212, as well as in GBK045827. The mechanism of resistance in FRAMIDA, IS9830, and SRN39 is reported to be low germination stimulation (Haussmann et al., 2000b; Mohamed et al., 2010; Gobena et al., 2017), which is the most widely studied mechanism of resistance to *Striga* in sorghum. The only other resistant check that clustered differently was N13, a durra sorghum from India, which is known to stimulate *Striga* germination but forms a mechanical barrier to *Striga* penetration (Maiti et al., 1984; Mohamed et al., 2016; Mbuvi et al., 2017). Genotype N13 grouped together with B35, a drought-tolerant variety, which has its origins in Ethiopia (Ochieng et al., 2020).

Screening for *Striga* Resistance in Pot and Field Trial Conditions

The trials made use of an existing *Striga*-infested field that was supplemented with artificial *Striga* inoculation, as well as a pot trial with artificial *Striga* inoculation to represent a second environment. One of the major challenges of undertaking reliable field trials using *Striga*-infested fields is the lack of homogeneity of *Striga* infections across different points of the field. This is especially due to the outcrossing nature of *Striga* and the tendency of its seeds to remain dormant in the soil for up to 20 years (Teka, 2014). In our investigation, both field trials and inoculated pots were used to take care of natural infection conditions, as well as enhance uniformity of infection. There are a number of successful *Striga* studies that used both field and pot experiments together with supplemented artificial inoculation in the same way (Kountche et al., 2013; Rodenburg et al., 2015; Midega et al., 2016; Abate et al., 2017). Other studies also made use of pot experiments as *Striga*-free control (Samejima et al., 2016), or to enable the isolation of root exudates (Jamil et al., 2011; Hooper et al., 2015). The low correlation reported here between field and pot experiments for *Striga*-related traits has

also been observed in other experiments (Haussmann et al., 2000a), suggesting that the use of pots for *Striga* screening should be completely discouraged.

Several factors could explain the observed differences in the pot and field experiments. First, the *Striga* infestation levels were significantly higher in the field trial than in the pot experiments, in which *Striga*-free soil was used before the addition of a standard amount of *Striga* in each pot. Second, it is likely that the field trial had a mixture of biotypes that had been accumulated over the years before the supplemented inoculation. Third, the *Striga* in the field trial was established exclusively under rainfed conditions, whereas the pot experiment was regularly watered to ensure sufficient moisture was maintained throughout the experiment. Furthermore, there could be rhizosphere differences that would affect the stimulation of *Striga* germination (Miché et al., 2000). Variable response to *Striga* under different test conditions has been observed in previous studies (Rao, 1984; Haussmann et al., 2000a).

Haussmann et al. (2004) hypothesized that abiotic stress, followed by ethylene production by microorganisms in the soil, could be responsible for the observed differences. In this study, drought stress could have been a factor in the field trial experiment, which was purely dependent on rain. Drought has been shown to induce strigolactone production in the roots (Haider et al., 2018), which in turn induces the germination of *Striga* (Cardoso et al., 1994). A significant decrease in rainfall in Kenya has been reported over the last four decades (Ayugi et al., 2016), suggesting that rainfed crops are highly likely to be exposed to drought during their growth periods. Rainfall data (not shown) from Alupe station during the growing season further confirmed the variability of rainfall that could have led to the exposure of the crops to drought. Nonetheless, we observed a number of genotypes showing consistency in their performance between the two trials for *Striga*-related traits.

Mechanism of Resistance to *Striga* in N13

Host plant resistance to *Striga* has been defined as the ability of the plant to reduce or prevent infection (Shew and Shew, 1994) through pre- and post-attachment mechanisms (Yoder and Scholes, 2010; Rodenburg et al., 2016). Tolerance refers to the extent to which effects of infection on the host plant are mitigated (Caldwell et al., 1958; Rodenburg et al., 2016). In this evaluation, N13 was not among the most resistant genotypes in both experiments. A study by Rodenburg et al. (2005) that incorporated N13, SRN39, IS9830, and FRAMIDA among other genotypes reported N13 as the most resistant genotype in the trial with ASNPC and NS_{max}-values significantly lower than all the other genotypes in the trial. Our results may suggest a possible breakdown of mechanical resistance that was initially recorded in N13, or potential contamination of seed source that may have led to the reduction in the resistance originally observed. However, these theories will need to be further investigated.

Although the genetic architecture of mechanical resistance to *Striga* in sorghum is yet to be understood, previous studies suggest that it is complex. QTL mapping studies in recombinant inbred populations using N13 as the resistant parent reported 11 and 9 QTLs in two different environments (Haussmann et al.,

2004). The durability of mechanical resistance will therefore depend on the factors at play, which may range from cell wall thickening, lignification, and silica deposition (Maiti et al., 1984). A recent study of *Striga* resistance in rice reported that enhanced lignin deposition and maintenance of the structural integrity of lignin polymers deposited at the infection site are crucial for post-attachment resistance against *S. hermonthica* (Mutuku et al., 2019). Similar studies will be necessary in sorghum to enhance our understanding of both pre- and post-attachment resistance to *Striga*. Such an outcome will enable the pyramiding of genes responsible for both pre- and post-attachment resistance in order to enhance the durability of resistance in the fields of farmers.

Drought Tolerance and Yield Stability Under *Striga* Infestation

A majority of the best yielding (most tolerant) genotypes were derived from crosses. However, the most consistent performance among these top-performing crosses was observed when crosses were made with any of the drought-tolerant genotypes, such as LODOKA, B35, and E36-1. B35 and E36-1 are drought-tolerant improved varieties that have been used for decades in the region and globally, while LODOKA is a drought-tolerant landrace. The superior and consistent agronomic performance of crosses involving drought-tolerant genotypes under *Striga* conditions was not surprising. Both drought stress and *S. hermonthica* infestation result in the production of abscisic acid (ABA) (Frost et al., 1997; Sah et al., 2016), which triggers stomatal closure (Kim et al., 2010). However, drought-tolerant sorghum genotypes have been shown to adapt to drought stress by preventing excessive ABA responses (Varoquaux et al., 2019). Such an adaptation response of drought-tolerant sorghum would also benefit their response to *S. hermonthica* and enhance the production of photosynthates to sustain plant growth and development in the presence of both stresses. A recent study in maize reported up to 19% increase in yield under stress in hybrids simultaneously expressing drought tolerance and *S. hermonthica* resistance as compared with those expressing only one of the traits (Menkir et al., 2020). Similar studies will need to be undertaken in sorghum to fully establish the interaction of drought and *S. hermonthica* stresses.

The older sources of drought tolerance will, however, need to be replaced with new superior performing genotypes such as F6YQ212 × B35 and ICSVIII IN × LODOKA identified in the current study. OKABIR, AKUOR-ACHOT, and LODOKA are landraces, which have also been recently reported to perform better than B35 and E36-1 under drought conditions (Ochieng et al., 2020). While LODOKA clustered with MACIA, both OKABIR and AKUOR-ACHOT appeared to be distantly related to MACIA, B35, and E36-1 and will therefore be good alternative sources of both drought and *Striga* tolerance. Better still, more focused introgression of *Striga* resistance and drought tolerance into farmer-preferred varieties should be planned to ensure better replacement of some of these old varieties.

Molecular Breeding for *Striga* Resistance and Tolerance

The demonstrated residual heterosis in HGW and NS_{max} at F_4 is great news for breeding programs as it shows the huge potential of enhancing the performance of varieties in response to *Striga* through improved genetics. Given the high variation in the ecotypes of *Striga* across different environments, the best breeding strategy would be genomic selection (GS) (Goddard, 2009). Our results provide a good basis for designing a GS strategy for developing *Striga*- and drought-tolerant sorghum varieties that will be suitable for the harsh environments typical of *Striga*-endemic ecologies. The available genomic resources in sorghum public databases will only enhance the ease with which GS is implemented in sorghum.

The DArT-seq technology (Sansaloni et al., 2011) that was used to characterize the germplasm proved to be a reliable and cost-effective technology for diversity analysis and confirming hybridity. While there are several studies reporting the use of DArT-seq for diversity analysis in sorghum (Kotla et al., 2019; Allan et al., 2020; Mengistu et al., 2020), this study is the first to use DArT-seq for hybridity testing. The unique SNP markers from this study will be useful for GS and for incorporation into marker panels that aim at the identification of successful hybrids from new crosses involving any of the parents. Future studies will also need to establish the specific molecular markers associated with new sources of resistance to *Striga* through genome-wide association studies (GWAS) or the characterization of biparental populations.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

KN and DO conceived the project, designed and supervised the experiment, and drafted the manuscript. NM did both the field and laboratory experiments and drafted the manuscript. EM provided some of the germplasm and supervised the fieldwork and data collection. WC supervised the fieldwork and data collection. MA supported the laboratory work and did part of the data analysis. DN and LW provided the genebank material and supervised the work, respectively. All authors approved the final version of the manuscript.

FUNDING

A postgraduate fellowship was provided to NM by the National Research Fund of Kenya (NRF) (<https://researchfund.go.ke/>). This study was partially funded by the CGIAR Research Program

(CRP) on Grain Legumes and Dryland Cereals. The wild accessions were collected as part of the initiative Adapting Agriculture to Climate Change: Collecting, Protecting and Preparing Crop Wild Relatives, which is supported by the Government of Norway. The project is managed by the Crop Trust with the Millennium Seed Bank of the Royal Botanic Gardens, Kew and implemented in partnership with national and international gene banks and plant breeding institutes around the world. For further information, see the project website: <http://www.cwrdiversity.org/>.

REFERENCES

- Aasha, M. B., Amani, E. H., Ras, H., Abusin, M. A., and Khalil, N. A. (2017). Effects of intercropping pearl millet with some legumes on *Striga hermonthica* emergence. *SSRG Int. J. Agric. Environ. Sci.* 4, 64–72. doi: 10.14445/23942568/IJAES-V4I6P113
- Abate, M., Hussien, T., Bayu, W., and Reda, F. (2017). Diversity in root traits of sorghum genotypes in response to *Striga hermonthica* infestation. *Weed Res.* 57, 303–313. doi: 10.1111/wre.12262
- Ali, R., Hash, C. T., Damaris, O., Elhussein, A., and Mohamed, A. H. (2016). Introgression of *striga* resistance into popular Sudanese sorghum varieties using marker assisted selection. *World J. Biotechnol.* 1, 48–55. doi: 10.1234/wjb.v1i1.34
- Aliche, E. B., Screpanti, C., Mesmaeker, A., Munnik, T., and Bouwmeester, H. J. (2020). Science and application of strigolactones. *New Phytol.* 227, 1001–1011. doi: 10.1111/nph.16489
- Allan, V., Vetriventhan, M., Senthil, R., Geetha, S., Deshpande, S., Rathore, A., et al. (2020). Genome-wide DArTSeq genotyping and phenotypic based assessment of within and among accessions diversity and effective sample size in the diverse sorghum, pearl millet, and pigeonpea landraces. *Front. Plant Sci.* 11:587426. doi: 10.3389/fpls.2020.587426
- Allard, R. W. (1960). *Principles of Plant Breeding*. New York, NY: John Wiley and Sons Inc.
- Ayugi, B., Wen, W., and Chepkemoi, D. (2016). Analysis of spatial and temporal patterns of rainfall variations over Kenya. *Environ. Earth Sci.* 6, 69–83.
- Beshir, M. M., Ahmed, N. E., Ali, A. M., Babiker, I. H., Rubaihayo, P., and Okori, P. (2015). Prevalence and severity of sorghum leaf blight in the sorghum growing areas of Central Sudan. *Wudpecker J. Agric. Res.* 4, 54–60. doi: 10.13140/RG.2.1.1785.1285
- Bouréma, D., Daouda, D., and James, H. W. (2005). Herbicide seed treatments for control of purple witchweed (*Striga hermonthica*) in sorghum and millet. *Weed Technol.* 19, 629–635.
- Bouwmeester, H., Schuurink, R. C., Bleeker, P. M., and Schiestl, F. (2019). The role of volatiles in plant communication. *Plant J.* 100, 892–907. doi: 10.1111/tjp.14496
- Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ramdoss, Y., and Buckler, E. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Caldwell, R. M., Schafer, J. F., Compton, L. E., and Patterson, F. L. (1958). Tolerance to cereal leaf rust. *Science* 128, 714–715
- Cardoso, C., Ruyter-Spira, C., Bouwmeester, H. J., Dawoud, D., and Sauerborn, J. (1994). Impact of drought stress and temperature on the parasitic weeds *Striga hermonthica* and *Alectra vogelii* in their early growth stages. *Exp. Agric.* 30, 249–257. doi: 10.1017/S0014479700024182
- Dembélé, B., Dembélé, D., and Westwood, J. (2005). Herbicide seed treatments for control of purple witchweed (*Striga hermonthica*) in sorghum and millet. *Weed Technol.* 19, 629–635. doi: 10.1614/WT-04-183R2.1
- Doggett, H. (1988). *Sorghum, 2nd Edn*. London: Longman, UK.
- Ejeta, G. (2007). Breeding for *Striga* resistance in sorghum: exploitation of intricate host parasite biology. *Crop Sci.* 47, S-216–S-227. doi: 10.2135/cropsci2007.04.0011IPBS
- Ejeta, G., and Knoll, J. E. (2007). “Marker-assisted selection in Sorghum,” in *Genomics-Assisted Crop Improvement: Vol 2: Genomics Applications in Crops*, eds R. K. Varshney and R. Tuberosa (Dordrecht: Springer), 187–205.
- Ezeaku, I. (2005). Development of sorghum populations for resistance to *Striga hermonthica* in the Nigerian Sudan Savanna. *African J. Biotechnol.* 3, 324–329. doi: 10.5897/AJB2004.000-2059
- Fantaye, B. (2018). Breeding sorghum for *Striga* resistance: a review. *J. Nat. Sci. Res.* 8, 1–8.
- FAOSTAT (2016). *Database of Agricultural Production FAO Statistical Databases (FAOSTAT)*. Available online at: <http://faostat.fao.org/default.aspx> (accessed April 26, 2021).
- Frost, D. L., Gurney, A. L., Press, M. C., and Scholes, J. D. (1997). *Striga hermonthica* reduces photosynthesis in sorghum: the importance of stomatal limitations and a potential role for ABA? *Plant Cell Environ.* 20, 483–492. doi: 10.1046/j.1365-3040.1997.d01-87.x
- Frost, H. M. (1994). *Striga Research and Survey in Kenya*. National Agricultural Research Project, KARI/ODA Crop Protection Project. Final Report. p. 61.
- Gobena, D., Shimelis, M., Rich, P. J., Ruyter-spira, C., Bouwmeester, H., and Kanuganti, S. (2017). Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes *Striga* resistance. *Proc. Natl. Acad. Sci. U.S.A.* 114, 4471–4476. doi: 10.1073/pnas.1618965114
- Goddard, M. (2009). Genomic selection: prediction of accuracy and maximisation of long-term response. *Genetica* 136, 245–257. doi: 10.1007/s10709-008-9308-0
- Gurney, A. L., Slate, J., Press, M. C., and Scholes, J. D. (2005). A novel form of resistance in rice to the angiosperm parasite *Striga hermonthica*. *New Phytol.* 169, 199–208. doi: 10.1111/j.1469-8137.2005.01560.x
- Haider, I., Andreo-Jimenez, B., Bruno, M., Bimbo, A., Floková, K., and Abuauf, H. (2018). The interaction of strigolactones with abscisic acid during the drought response in rice. *J. Exp. Bot.* 69, 2403–2414. doi: 10.1093/jxb/ery089
- Hallauer, A. R., Carena, M. J., and Miranda Filho, J. B. (2010). *Quantitative Genetics in Maize Breeding*. New York, NY: Springer.
- Hausmann, B. I. G., Hess, D. E., Omany, G. O., Folkertsma, R. T., Reddy, B. V., Kayentao, M., et al. (2004). Genomic regions influencing resistance to the parasitic weed *Striga hermonthica* in two recombinant inbred populations of sorghum. *Theor. Appl. Genet.* 109, 1005–1016. doi: 10.1007/s00122-004-1706-9
- Hausmann, B. I. G., Hess, D. E., Reddy, B. V. S., Mukuru, S. Z., Kayentao, M., Welz, H. G., et al. (2000a). Quantitative-genetic parameters of sorghum growth under *Striga* infestation in Mali and Kenya. *Plant Breed.* 120, 49–56. doi: 10.1046/j.1439-0523.2001.00546.x
- Hausmann, B. I. G., Hess, D. E., Reddy, B. V. S., Welz, G., and Geiger, H. H. (2012). Analysis of resistance to *Striga hermonthica* in diallel crosses of sorghum. *Euphytica* 116, 33–40. doi: 10.1023/A:1004046001009
- Hausmann, B. I. G., Hess, D. E., Welz, H. G., and Geiger, H. H. (2000b). Improved methodologies for breeding *Striga*-resistant sorghums. *Field Crops Res.* 66, 195–211. doi: 10.1016/S0378-4290(00)00076-9
- Hess, D. E., and Ejeta, G. (1992). Inheritance of Resistance to *Striga* in Sorghum Genotype SRN39. *Plant Breed.* 109, 233–241. doi: 10.1111/j.1439-0523.1992.tb00178.x
- Hooper, A. M., Caulfield, J. C., Hao, B., Pickett, J. A., Midega, C. A. O., and Khan, Z. R. (2015). Isolation and identification of Desmodium root exudates from drought tolerant species used as intercrops against *Striga hermonthica*. *Phytochemistry* 117, 380–387. doi: 10.1016/j.phytochem.2015.06.026

ACKNOWLEDGMENTS

The authors are grateful to the University of Nairobi for facilitation, technical support, and knowledge sharing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.671984/full#supplementary-material>

- Jamil, M., Charnikhova, T., Cardoso, C., Jamil, T., Ueno, K., Verstappen, F., et al. (2011). Quantification of the relationship between strigolactones and *Striga hermonthica* infection in rice under varying levels of nitrogen and phosphorus. *Weed Res.* 51, 373–385. doi: 10.1111/j.1365-3180.2011.00847.x
- Khan, Z., Midega, C., Pittchar, J., Pickett, J., and Bruce, T. (2011). Push–pull technology: a conservation agriculture approach for integrated management of insect pests, weeds, and soil health in Africa. *Int. J. Agric. Sustain.* 9, 162–170. doi: 10.3763/ijas.2010.0558
- Khan, Z. R., Pickett, J. A., Wadhams, L. J., Hassanali, A., and Midega, C. A. O. (2006). Combined control of *Striga hermonthica* and stem borers by maize–*Desmodium* spp. intercrops. *Crop Prot.* 25, 989–995. doi: 10.1016/j.cropro.2006.01.008
- Kidanemariam, W., Kassahun, B., and Taye, T. (2018). *Assessment of Heterotic Performance and Combining Ability of Ethiopian Elite Sorghum (Sorghum bicolor (L.) Moench) Lines*. Masters Thesis, Jimma University, Ethiopia.
- Kim, S. K., Adetimirin, V. O., Thé, C., and Dossou, R. (2002). Yield losses in maize due to *Striga hermonthica* in West and Central Africa. *Int. J. Pest Manag.* 48, 211–217. doi: 10.1080/09670870110117408
- Kim, T. H., Bohmer, M., Hu, H., Nishimura, N., and Schroeder, J. I. (2010). Guard cell signal transduction network: advances in understanding abscisic acid, CO₂, and Ca²⁺ signaling. *Annual Rev. Plant Biol.* 61, 561–591. doi: 10.1146/annurev-arplant-042809-112226
- Kiprotich, F., Mwendia, M. C., Cheruiyot, K. E., and Wachira, N. F. (2015). Nutritional suitability of bred sorghum (*Sorghum bicolor*) accessions from East Africa. *African J. Food Sci.* 9, 326–334. doi: 10.5897/AJFS2015.1288
- Kotla, K., Phuke, R., Hariprasanna, K., Mehre, S. P., Rathore, A., Gorthy, S., et al. (2019). Identification of QTLs and candidate genes for high grain Fe and Zn concentration in sorghum [*Sorghum bicolor* (L.) Moench]. *J. Cereal Sci.* 90:102850. doi: 10.1016/j.jcs.2019.102850
- Kountche, B. A., Hash, T. C., Dodo, H., Laoualy, O., Sanogo, M. D., Timbeli, A., et al. (2013). Development of a pearl millet *Striga*-resistant gene pool: Response to five cycles of recurrent selection under *Striga*-infested field conditions in West Africa. *Field Crops Res.* 154, 82–90. doi: 10.1016/j.fcr.2013.07.008
- MacOpio, L., Vitale, J., and Sanders, J. (2010). *An Ex-Ante Assessment of a Striga Control Programme in East Africa*. Kilimo Trust, p. 6–25. Available online at: <http://www.parasite-project.org/wp-content/uploads/2014/02/Kilimo-trust-2009-Ex> (accessed January 26, 2021).
- Maiti, R. K., Ramaiah, K. V., Bisen, S. S., and Chidley, V. L. (1984). A comparative study of the haustorial development of *Striga asiatica* (L.) Kuntze on Sorghum cultivars. *Ann. Bot.* 54, 447–457. doi: 10.1093/oxfordjournals.aob.a086816
- Mallu, T. S., Mutinda, S., Githiri, S. M., Odeny, D., and Runo, S. (2021). New pre-attachment *Striga* resistant sorghum adapted to African agro-ecologies. *Pest Manag. Sci.* 77, 2894–2902. doi: 10.1002/ps.6325
- Marley, P. S., Diourte, M., Neya, A., and Rattunde, F. W. (2005). Sorghum anthracnose and sustainable management strategies in West and Central Africa. *J. Sustain. Agric.* 25, 43–56. doi: 10.1300/J064v25n01_05
- Mbuvi, D. A., Masiga, C. W., Kuria, E., Masanga, J., Wamalwa, M., Mohamed, A., et al. (2017). Novel sources of witchweed (*Striga*) resistance from wild sorghum accessions. *Front. Plant Sci.* 8:116. doi: 10.3389/fpls.2017.00116
- Mengistu, D., Shimelis, H., Laing, M., Lule, D., Assefa, E., and Mathew, I. (2020). Genetic diversity assessment of sorghum (*Sorghum bicolor* (L.) Moench) landraces using SNP markers. *South African J. Plant Soil* 37, 220–226. doi: 10.1080/02571862.2020.1736346
- Menkir, A., Adetimirin, V. O., Yallou, C. G., and Gedil, M. (2010). Relationship of genetic diversity of inbred lines with different reactions to *Striga hermonthica* (Del.) benth and the performance of their crosses. *Crop Sci.* 50, 602–611. doi: 10.2135/cropsci2009.05.0247
- Menkir, A., Crossa, J., Meseke, S., Bossey, B., Muhyideen, O., Riberio, P. F., et al. (2020). Staking tolerance to drought and resistance to a parasitic weed in tropical hybrid maize for enhancing resilience to stress combinations. *Front. Plant Sci.* 11:166. doi: 10.3389/fpls.2020.00166
- Miché, L., Bouillant, M. L., Rohr, R., Sallé, G., and Bally, B. (2000). Physiological and cytological studies on the inhibition of striga seed germination by the plant growth-promoting bacterium *Azospirillum brasilense*. *Eur. J. Plant Pathol.* 106, 347–351. doi: 10.1023/A:1008734609069
- Midega, C. A., Pickett, J., Hooper, A., Pittchar, J., and Khan, Z. R. (2016). Maize landraces are less affected by *Striga hermonthica* relative to hybrids in western Kenya. *Weed Technol.* 30, 21–28. doi: 10.1614/WT-D-15-00055.1
- Mohamed, A. H., Housley, T. L., and Ejeta, G. (2010). An in vitro technique for studying specific *Striga* resistance mechanisms in sorghum. *Afr. J. Agric. Res.* 5, 1868–1875. doi: 10.5897/AJAR.9000044
- Mohamed, N., Charnikhova, T., Bakker, E. J., van Ast, A., Babiker, A. G., and Bouwmeester, H. J. (2016). Evaluation of field resistance to *Striga hermonthica* (Del.) Benth. in *Sorghum bicolor* (L.) Moench. The relationship with strigolactones. *Pest Manag. Sci.* 72, 2082–2090. doi: 10.1002/ps.4426
- Mutuku, J. M., Cui, S., Hori, C., Takeda, Y., Tobimatsu, Y., Nakabayashi, R., et al. (2019). The structural integrity of lignin is crucial for resistance against *Striga hermonthica* parasitism in rice. *Plant Physiol.* 179, 1796–1809. doi: 10.1104/pp.18.01133
- Mwenda, E. T. (2019). *Phenotypic and Biochemical Screening of Sorghum Genotypes for Growth and Rice Weevil Resistance in Tanzania*, Doctoral dissertation, Nelson Mandela African Institution of Science and Technology (NM-AIST), Tanzania.
- Ngugi, K., Ngugi, A. J., Osama, S., and Mugoya, C. (2015). Combating *Striga* weed in sorghum by transferring resistance quantitative trait loci through molecular marker assisted introgression. *J. Plant Breed. Genet.* 3, 67–76.
- Ochieng, G., Ngugi, K., Wamalwa, L. N., Manyasa, E., Muchira, N., Nyamongo, D. A., et al. (2020). Novel sources of drought tolerance from landraces and wild sorghum relatives. *Crop Sci.* 61, 104–118. doi: 10.1002/csc2.20300
- Ohadi, S., Hodnett, G., Rooney, W., and Bagavathiannan, M. (2017). Gene flow and its consequences in Sorghum spp. *Crit. Rev. Plant Sci.* 36, 5–6. doi: 10.1080/07352689.2018.1446813
- Oswald, A., and Ransom, J. (2001). *Striga* control and improved farm productivity using crop rotation. *Crop Prot.* 20, 113–120. doi: 10.1016/S0261-2194(00)00063-6
- Pamer, E. G. (2016). Resurrecting the intestinal microbiota to combat antibiotic-resistant pathogens. *Science* 352, 535–538. doi: 10.1126/science.aad9382
- Pieterse, A. H., and Pesch, C. J. (1983). The witch weeds (*Striga* spp.) — a review. *Abstr. Trop. Agric.* 9:8.
- Rao, M. J. V. (1984). “Patterns of resistance to *Striga asiatica* in sorghum and millets, with special reference to Asia,” in *Workshop on the Biology and Control of Striga*, 14–17 November 1983, (Dakar).
- Rebeka, G., Hussein, S., Mark, L., Pangirayi, T., and Mandefro, N. (2013). Evaluation of sorghum genotypes compatibility with *Fusarium oxysporum* under *Striga* infestation. *Crop Sci.* 53, 385–393. doi: 10.2135/cropsci2012.02.0101
- Rich, P. J., Grenier, C., and Ejeta, G. (2004). *Striga* resistance in the wild relatives of sorghum. *Crop Sci.* 44, 2221–2229. doi: 10.2135/cropsci2004.2221
- Robinson, H. F., Comstock, R. E., and Harvey, P. H. (1949). Estimates of heritability and degree of dominance in corn. *Agron. J.* 41, 353–359. doi: 10.2134/agronj1949.00021962004100080005x
- Rodenburg, J., Bastiaans, L., and Kropff, M. J. (2006). Characterization of host tolerance to *Striga hermonthica*. *Euphytica* 147, 353–365. doi: 10.1007/s10681-005-9030-2
- Rodenburg, J., Bastiaans, L., Weltzien, E., and Hess, D. E. (2005). How can field selection for *Striga* resistance and tolerance in sorghum be improved? *Field Crops Res.* 93, 34–50. doi: 10.1016/j.fcr.2004.09.004
- Rodenburg, J., Cissoko, M., Kaye, J., Dieng, I., Khan, Z. R., Midega, C. A. O., et al. (2015). Do NERICA rice cultivars express resistance to *Striga hermonthica* (Del.) Benth. and *Striga asiatica* (L.) Kuntze under field conditions? *Field Crops Res.* 170, 83–94. doi: 10.1016/j.fcr.2014.10.010
- Rodenburg, J., Demont, M., Zwart, S. J., and Bastiaans, L. (2016). Parasitic weed incidence and related economic losses in rice in Africa. *Agric. Ecosyst. Environ.* 235, 306–317. doi: 10.1016/j.agee.2016.10.020
- Rubiales, D., and Fernández-Aparicio, M. (2012). Innovations in parasitic weeds management in legume crops. A review. *Agron. Sustain. Dev.* 32, 433–449. doi: 10.1007/s13593-011-0045-x
- Sah, S. K., Reddy, K. R., and Li, J. (2016). Abscisic acid and abiotic stress tolerance in crop plants. *Front. Plant Sci.* 7:571. doi: 10.3389/fpls.2016.00571
- Samejima, H., Babiker, A. G., Takikawa, H., Sasaki, M., and Sugimoto, Y. (2016). Practicality of the suicidal germination approach for controlling *Striga hermonthica*. *Pest Manag. Sci.* 72, 2035–2042. doi: 10.1002/ps.4215
- Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., et al. (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative

- genotyping for molecular breeding of Eucalyptus. *BMC Proc.* 5:P54. doi: 10.1186/1753-6561-5-S7-P54
- Satish, K., Gutema, Z., Grenier, C., Rich, P. J., and Ejeta, G. (2012). Molecular tagging and validation of microsatellite markers linked to the low germination stimulant gene (*lgs*) for *Striga* resistance in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor. Appl. Genet.* 124, 989–1003. doi: 10.1007/s00122-011-1763-9
- Shew, H. D., and Shew, B. B. (1994). "Host resistance," in *Epidemiology and Management of Root Diseases*, eds C. L. Campbell and D. M. Benson (Berlin: Springer Verlag), 244–275.
- Syukur, M., Sujiprihati, S., and Yudianti, R. (2012). *Teknik and Pemuliaan Tanaman*. Jakarta: Penebar Swadaya.
- Teka, H. B. (2014). Advance research on *Striga* control: a review. *African J. Plant Sci.* 8, 492–506. doi: 10.5897/AJPS2014.1186
- Tuinstra, M., Soumana, S., Al-Khatib, K., Kapran, I., Toure, A., Ast, A., et al. (2009). Efficacy of Herbicide Seed Treatments for Controlling *Striga* Infestation of Sorghum. *Crop Sci.* 49, 923–921. doi: 10.2135/cropsci2008.06.0357
- Turner, J. H. (1953). A study of heterosis in upland cotton II. Combining ability and inbreeding effects 1. *Agron. J.* 45, 487–490. doi: 10.2134/agronj1953.00021962004500100008x
- Van Delft, G. J., Graves, J. D., Fitter, A. H., and Pruiksma, M. (1997). Spatial distribution and population dynamics of *Striga hermonthica* seeds in naturally infested farm soils. *Plant Soil* 195, 1–15. doi: 10.1023/A:1004214015281
- Van Hast, A., Bastiaans, L., and Kropff, M. J. (2000). A comparative study on *Striga hermonthica* interaction with a sensitive and a tolerant sorghum cultivar. *Weed Res.* 40, 479–493. doi: 10.1046/j.1365-3180.2000.00204.x
- Varoquaux, N., Cole, B., Gao, C., Pierroz, G., Baker, C. R., Patel, D., et al. (2019). Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses. *Proc. Natl. Acad. Sci. U.S.A.* 116, 27124–27132. doi: 10.1073/pnas.1907500116
- VSN International. (2017). *Genstat for Windows*, 19th edition. UK: Wang: Hemel Hempstead.
- Yoder, J. I., and Scholes, J. D. (2010). Host plant resistance to parasitic weeds; recent progress and bottlenecks. *Curr. Opin. Plant Biol.* 13, 478–484. doi: 10.1016/j.pbi.2010.04.011
- Yohannes, T., Abraha, T., Kiambi, D., Folkertsma, R., Tom Hash, C., Ngugi, K., et al. (2015). Marker-assisted introgression improves *Striga* resistance in an Eritrean farmer-preferred sorghum variety. *Field Crops Res.* 173, 22–29. doi: 10.1016/j.fcr.2014.12.008
- Yoshida, S., and Shirasu, K. (2009). Multiple layers of incompatibility to the parasitic witchweed, *Striga hermonthica*. *New Phytol.* 183, 180–189. doi: 10.1111/j.1469-8137.2009.02840.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Muchira, Ngugi, Wamalwa, Avosa, Chepkorir, Manyasa, Nyamongo and Odeny. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Differential Modulation of Heat-Inducible Genes Across Diverse Genotypes and Molecular Cloning of a sHSP From Pearl Millet [*Pennisetum glaucum* (L.) R. Br.]

S. Mukesh Sankar^{1*}, C. Tara Satyavathi^{2*}, Sharmistha Barthakur³, Sumer Pal Singh¹, C. Bharadwaj¹ and S. L. Soumya¹

¹ Division of Genetics, Indian Council of Agricultural Research-Indian Agricultural Research Institute, New Delhi, India, ² Indian Council of Agricultural Research-All India Coordinated Research Project on Pearl Millet (AICPMIP), Jodhpur, India, ³ Functional Genomics, Indian Council of Agricultural Research-National Institute for Plant Biotechnology (NIPB), New Delhi, India

OPEN ACCESS

Edited by:

Mahalingam Govindaraj,
HarvestPlus–Alliance
Bioversity-CIAT, Colombia

Reviewed by:

Sudhakar Reddy Palakolanu,
International Crops Research Institute
for the Semi-Arid Tropics, Niger
Prachi Pandey,
National Institute of Plant Genome
Research (NIPGR), India

*Correspondence:

S. Mukesh Sankar
mukesh.genetics@gmail.com
C. Tara Satyavathi
csatyavathi@gmail.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 28 January 2021

Accepted: 08 June 2021

Published: 16 July 2021

Citation:

Mukesh Sankar S, Tara Satyavathi C,
Barthakur S, Singh SP, Bharadwaj C
and Soumya SL (2021) Differential
Modulation of Heat-Inducible Genes
Across Diverse Genotypes and
Molecular Cloning of a sHSP From
Pearl Millet [*Pennisetum glaucum* (L.)
R. Br.]. *Front. Plant Sci.* 12:659893.
doi: 10.3389/fpls.2021.659893

The survival, biomass, and grain yield of most of the crops are negatively influenced by several environmental stresses. The present study was carried out by using transcript expression profiling for functionally clarifying the role of genes belonging to a small heat shock protein (sHSP) family in pearl millet under high-temperature stress. Transcript expression profiling of two high-temperature-responsive marker genes, *Pgcp70* and *PgHSF*, along with physio-biochemical traits was considered to screen out the best contrasting genotypes among the eight different pearl millet inbred lines in the seedling stage. Transcript expression pattern suggested the existence of differential response among different genotypes upon heat stress in the form of accumulation of heat shock-responsive gene transcripts. Genotypes, such as WGI 126, TT-1, TT-6, and MS 841B, responded positively toward high-temperature stress for the transcript accumulation of both *Pgcp70* and *PgHSF* and also indicated a better growth under heat stress. PPML-69 showed the least responsiveness to transcript induction; moreover, it supports the membrane stability index (MSI) data for scoring thermotolerance, thereby suggesting the efficacy of transcript expression profiling as a molecular-based screening technique for the identification of thermotolerant genes and genotypes at particular crop growth stages. The contrasting genotypes, such as PPML-69 (thermosusceptible) and WGI-126 and TT-1 (thermotolerant), are further utilized for the characterization of thermotolerance behavior of sHSP by cloning a *PgHSP16.97* from the thermotolerant cv. WGI-126. In addition, the investigation was extended for the identification and characterization of 28 different HSP20 genes through a genome-wide search in the pearl millet genome and an understanding of their expression pattern using the RNA-sequencing (RNA-Seq) data set. The outcome of the present study indicated that transcript profiling can be a very useful technique for high-throughput screening of heat-tolerant genotypes in the seedling stage. Also, the identified *PgHSP20s* genes can provide further insights into the molecular regulation of pearl millet stress tolerance, thereby bridging them together to fight against the unpredicted nature of abiotic stress.

Keywords: Thermotolerance, Heat Shock Proteins, alpha-crystalline domain, Pearl Millet, Cloning, qRT-PCR, RNA-seq

INTRODUCTION

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is an annual warm season, C₄ cereal crop that is cultivated extensively in the Indian subcontinent and African semiarid regions. It covers an area of 29 million ha and forms the staple food and fodder for 90 million resource-poor inhabitants (Sharma et al., 2021). With a total area of 7.41 million ha, India is the largest producer of pearl millet in the world, having an annual production of 10.3 million tons in 2020–2021 (Indiastat, 2020). Pearl millet is mostly grown as a *Kharif* crop with the least input during the hottest periods by the onset of monsoon. A high temperature of more than 42°C accompanied by moisture stress during the seedling stage of the crop, particularly at germination and seedling establishment, affects the plant status and thereby its productivity. A rise in air and soil temperature of the semiarid regions of India and sub-Saharan Africa due to global climate change continues to be a severe constraint in the growth and development of the crop; thereby causing a huge loss in terms of quantity and quality (Ullah et al., 2018). Such a situation demands critical attention not only for the development of stress-tolerant genotypes but also for the identification and characterization of genes responsible for thermotolerance.

Plants being sessile can dramatically alter their gene expressions in response to various stress signals through a set of morphological, physiological, and molecular changes that can harm the stability of plant growth and productivity (Castelán-Muñoz et al., 2019). The activation and regulation of stress-related genes play a significant role in the acquisition of thermotolerance among plants. In response to heat stress, pearl millet produces an array of proteins that help in alleviating the effects of stress. One such major protein family is heat shock proteins (HSPs). Both stress signal transduction and gene activation are aided by HSPs/chaperones (Nievolá et al., 2017). Heat stress-responsive signal transduction pathways and defense mechanisms involve heat shock factors (HSFs) and HSPs in sensing reactive oxygen species (ROS) produced under heat stress and in preventing protein misfolding (Sajid et al., 2018). For the first time, Howarth. C. J. (1991) reported that specific HSP transcripts aid in the diagnosis of plant stress and play a role in the acquisition of thermotolerance in pearl millet. Based on their approximate molecular weights, five major families of HSPs, namely HSP100, HSP90, HSP70, HSP60, and the HSP20 or small HSP (sHSP), are recognized (Waters and Vierling, 2020). HSP20s are the diverse groups of proteins that are preserved in both eukaryotes and prokaryotes with molecular weights in the range of 11–35 kDa, the expression of which is limited in the absence of environmental stress but upregulated to over 200-fold upon the induction of heat stress (Wahid et al., 2007). By protecting the native proteins from irreversible aggregation and oxidative inactivation, it plays a decisive role in the defense of an organism during physiological stress (Ma et al., 2021). One of the research findings further showed that sHSPs, along with HSP70, play a key role in quality control, possibly contributing to the maintenance of cell membrane integrity, especially under stress (Savic et al., 2012). Hence, the maintenance of cell membrane fluidity under high-temperature stress is due to a higher expression of these two

major genes, its expression pattern across genotypes can be used for screening seedling thermotolerance.

HSP20 gene families have been reported in *Arabidopsis*, rice, tomato, potato, watermelon, apple, grapes, and some other plant species (Scharf et al., 2001; Ouyang et al., 2009; Sarkar et al., 2009; Yu et al., 2016; Zhao et al., 2018; He et al., 2019; Ji et al., 2019; Yao et al., 2020). Furthermore, HSP20 genes are organized into multigenic families, with defined classification and biological functions (Waters et al., 1996; Scharf et al., 2001; Siddique et al., 2008; Sarkar et al., 2009; Zhang et al., 2016). Although the available pearl millet whole-genome sequence is a vital asset for an in-depth understanding of various gene families distributed among the genomes (Varshney et al., 2017), little knowledge on the integrated HSP20 family at the whole genomic level in pearl millet exists. The research findings will assist the researcher to truly assess the functional role of the pearl millet HSP20 gene family and its role in stress adaptation. These genes can pose a great opportunity in the development of a thermotolerant cultivar in pearl millet and other cereals.

In conventional breeding for heat tolerance, the identification of reliable and effective screening methods to facilitate the identification of heat-tolerant plants and the genes responsible for thermotolerance represents a major challenge. The transcript expression method based on semi-quantitative reverse transcription- (RT-) PCR offers a rapid and cost-effective way to screen out large germplasm lines or advance breeding lines that were inferior under specific thermal regimes to more accurate methods such as quantitative real-time PCR (qRT-PCR) or RNA-sequencing (RNA-Seq). Recently, various researchers employed semi-quantitative RT-PCR for the successful screening of germplasm and for fishing out the candidate genes related to drought and heat stresses (Bharadwaj et al., 2019; Choi et al., 2019; Galvez et al., 2020).

In this context, the current study aimed (1) to screen the most contrasting genotypes against heat stress using the physiological and biochemical parameters, (2) to confirm the results of the abovementioned screening techniques with the semi-quantitative RT-PCR expression profile of key heat-responsive marker genes, such as *Pgcp70* and *PgHSF*, (3) to identify, clone, and characterize a *PgHSP20* gene from the most thermotolerant genotype, and (4) to distinguish various members of the pearl millet *HSP20* gene family using a bioinformatics method and to characterize the integration of sequence features, chromosome location, phylogenetic relationship, and expression patterns. This study unravels the effectiveness of reverse transcription PCR- (RT-PCR) based screening methods for the identification and utilization of the thermotolerant genes from the superior heat-tolerant genotypes to bridge the supra-optimal temperature tolerance with high productivity in pearl millet.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

A total of eight elite pearl millet inbred lines (six thermotolerant and two thermosusceptible), based on our earlier studies

(Mukesh Sankar et al., 2014; James et al., 2015), were collected from the pearl millet breeding unit of the Indian Agricultural Research Institute (IARI), New Delhi, India. The pedigree, salient characteristics, and breeding station from where it has been developed are presented in **Table 1**. The experiments were done in a completely randomized block design with three replications. Seeds were surface sterilized and grown in a pot containing autoclaved Soilrite™ and were kept with a 16-h photoperiod at 25°C in a culture room as control until subjecting them to heat stress treatment in a growth chamber (**Supplementary Figure 1**).

Screening of Genotypes Using Physiological and Biochemical Parameters

Preliminary analysis of the already known physiological and biochemical marker traits for seedling thermotolerance was undertaken to determine the most thermotolerant and thermosusceptible ones among the eight genotypes. Hence, 10-day-old seedlings were subjected to heat stress at 42°C for 6 h.

The membrane stability index (MSI) of both control and stressed seedlings was assessed by using the modified method given by Sairam and Tyagi (2004) with the formula:

$$\text{MSI (\%)} = [1 - (C_1/C_2)] \times 100$$

Relative water content (RWC) of both control and stressed seedlings was estimated by applying the formula:

$$\text{RWC (\%)} = [(FW - DW)/(TW - DW)] \times 100$$

as described by Yang et al. (2011).

Leaf absorbance was measured with a chlorophyll meter (SPAD-502, Minolta, Tokyo, Japan) as a non-destructive tool for estimating leaf chlorophyll. The lipid peroxidation level for all the samples was estimated through the thiobarbituric acid-reactive substance (TBARS) assay proposed by Heath and Packer (1968).

Heat Stress Treatment, RNA Isolation, and Transcript Expression Profiling

Heat stress was imposed on 7- and 10-day-old seedlings in a controlled growth chamber at 42°C for 2 and 6 h, respectively. The aerial portion of seedlings was harvested after heat treatment given in a growth chamber. Simultaneously, another set of plant samples kept as control were also collected under controlled conditions. To preserve the stage-specific transcript, all collected tissue samples were wrapped in an aluminum foil, labeled, and flash-frozen by using liquid nitrogen before being stored at −80°C for RNA isolation. Total RNA was isolated from the tissues using the TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The isolated RNAs were checked in 1.2% agarose gel. About 1 µg of the samples were then treated with DNase (Promega Corporation, Madison, WI, USA) and subjected to NanoDrop spectrophotometer quantification (ND-1000, Thermo Fisher Scientific, Waltham, MA, USA). DNase-treated RNA was reverse transcribed by using the Thermo, Verso™ cDNA-kit (Thermo Fisher Scientific Inc., Waltham, MA, USA). For each reaction, 50 ng of complementary DNA (cDNA) was used as a template, accompanied by the amplification with gene-specific primers.

The transcript expression patterns of two well-known pearl millet genes, namely HSP70 and HSF, were compared to identify the best thermotolerant and thermosusceptible ones among the selected pearl millet genotypes in response to heat stress. The Pearson correlation coefficient was also estimated to conclude the relationship among the studied physiological and biochemical traits with the gene expression level of the candidate genes. Coding sequences of these two representative pearl millet genes, i.e., *Pgcp70* (Acc. no. EF495353.1) and *Pghsf* (Acc. no. EU492460.1), were downloaded from NCBI (<http://www.ncbi.nlm.nih.gov>). Specific primers were designed to amplify the selected messenger RNA (mRNA) sequence of the abovementioned genes for the semiquantitative RT-PCR expression analysis while maintaining stringency and specificity. The details of the primers along with their melting temperatures (T_m) are presented in **Table 2**. RT-PCR reactions were then performed by using the PCR master mix (Promega Corporation,

TABLE 1 | Details of genotypes used for transcript expression studies for heat tolerance in pearl millet.

Sl. No	Genotype	Codes	Pedigree, salient characteristics, and breeding use	Seed source ^a	Mean MSI (%) ^b
1	MS 411 B	G1	Extra early male sterile line developed from 263 B through selection.	IARI	56.38
2	MS 841 B	G2	DM-resistant selection from residual variability in 5141B; seed parent of Pusa 23, Pusa 322, and Pusa 605	IARI-ICRISAT	71.21
3	D-23	G3	DM-resistant derivative of K-560-230; Male parent of Pusa 23	IARI	65.84
4	PPMI-69	G4	Derivative of PPMI 43; Male parent of Pusa 605	IARI	52.53
5	PPMI-301	G5	Derivative of a cross between four elite restorers having bold ear head; Male parent of Pusa 322	IARI	67.29
6	TT-1	G6	Selection from line no. 868 suited for arid regions of Jodhpur	CAZRI	67.99
7	TT-6	G7	Selection from line no. 873 suited for arid regions of Jodhpur	CAZRI	70.65
8	WGI-126	G8	Pearly white, bold grained inbred with sturdy stem and broad leaf	IARI	67.65

^aCAZRI, Central Arid Zone Research Institute, Jodhpur (India); ICRISAT, International Crop Research Institute for SemiArid Tropics, Hyderabad (India); IARI, Indian Agricultural Research Institute, New Delhi (India).

^bMean MSI value.

Source: Mukesh Sankar et al. (2014) (MSI > 65% were thermotolerant).

Madison, WI, USA) as per the instruction of the manufacturer in a volume of 50- μ l reaction mixture. A total of 30 cycles of PCR with 4 min of an initial denaturation at 94°C, 94°C for 45 s, 48–60°C (T_m optimized for the individual gene) for 45 s, 72°C for 1-min amplification with a final extension at 72°C for 10 min were performed. The RT-PCR products were loaded on 1.2% agarose gel, and the stained DNA products were photographed by using Alfa Imager gel documentation system HP (Protein Simple, San Jose, CA, USA). In all expression studies, the housekeeping gene actin was used as a reference gene (internal constitutive control) to show equal loading of cDNA and to ensure its integrity under various degrees of heat stress. The transcript levels of each target gene were averaged over three reactions. The gene expression data were normalized by deducting the mean expression level from the reference gene. The relative fold change in the expression of treatments (T) was compared with those from the regular growth stage (C) and done by the expression value of the susceptible genotype control as a calibrator for the respective genotype using Alfa Imager Software tools by keeping the density of bands in the control as unity.

Isolation and Cloning of a Full-Length *PgHSP16.97*

The full-length ORF of *PgHSP16.97* was amplified from the pearl millet inbred genotype, WGI 126, by using the primers designed based on the sequence homology with other grasses sequences available in the public domain (Table 2). The sequence was cloned in a pGMT vector, confirmed by colony PCR and restriction digestion followed by Sanger sequencing (Xcelris Labs Ltd., Ahmedabad, India). The sequence homology was carried out by using NCBI (www.ncbi.nlm.nih.gov) tools and verified and deposited in GenBank. The conceptual

translation of nucleotide sequence was undertaken by using the ExPasy translate tool (<http://web.expasy.org/translate/>). Multiple sequence alignment were carried out by using the CLUSTALW software package (Thompson et al., 1994) with full-length HSP 17.0 protein sequences publicly available for *P. glaucum*, *Zea mays*, *Oryza sativa*, *Triticum aestivum*, *Sorghum bicolor*, and *Hordeum* spp.

Expression Analysis of *PgHSP16.97* in Different Tissues at Different Developmental Stages

To investigate the expression pattern of HSP 16.97, the tissue samples from the whole seedling at 10 days after sowing (DAS), leaves, stems, and roots at 30 DAS, and flag leaf and developing panicle at 55 DAS were also collected from both control and heat-stressed plant at 42°C for 6 h. Semi-quantitative RT-PCR was carried out by using the same gene-specific primers for the cloning of the gene for an initial analysis. The GenScript Real-time PCR Primer design tool (<https://www.genscript.com/ssl-bin/app/primer>) was used to design the qRT-PCR primers (Table 3). The SYBR Green Master Mix (2X) was utilized according to the instruction of the manufacturer to study the gene expression (Thermo Fisher Scientific, Waltham, MA, USA). qRT-PCR analysis was performed in three technical replicates (obtained by pooling the tissues from the three biological samples of each treatment) and with the following thermal cycles: 1 cycle at 95°C for 30 s, followed by 40 cycles alternatively at 95°C for 5 s, and 58°C for 1 min in the StepOne Real-Time PCR system (Applied Biosystems, Waltham, MA, USA). Based on the comparative Ct method using the $2^{-\Delta\Delta C_T}$ method, the relative expression levels of the *PgHSP16.97* gene were estimated (Livak and Schmittgen, 2001). A constitutive *PgAct* gene-based

TABLE 2 | Details of primers used for transcript expression profiling and full-length cloning by reverse transcription PCR (RT-PCR).

Gene name	Strand	Primer Sequence (5'→3')	Expected product size	Tm
<i>Pgcp70</i>	F	5' ACAGGGAAGAAGCAGGACATGACA 3'	184 bp	50.0°C
	R	5' AGCTCCTTGAGTTGCTTCTCGGTT 3'		
<i>PgHSF</i>	F	5' ATATCTTCGCCTCCCTCAGGGTGATA 3'	145 bp	48.0°C
	R	5' GTATGAAGGCAACACACCACGCAA 3'		
<i>Pg actin</i>	F	5' AGCGAGTCTTCATAGGGCGATTGT 3'	200 bp	60.0°C
	R	5' TAGCTCTGGGTTTCGAGTGGCATT 3'		
<i>PgHSP16.97</i>	F	5' AGTTTCAGCAATGTCGCTGGT 3'	560 bp	53.0°C
	R	5' ACAAGCACGACTCGTAGCATC 3'		

Where F, indicates the forward primer and R, indicates the reverse primer.

TABLE 3 | Details of gene-specific primers for qRT-PCR gene expression analysis.

Gene name	Strand	Primer sequence (5'→3')	Primer Tm (°C)	Primer length (bp)	Amplicon size (bp)
<i>PgHSP16.97RT</i>	F	CAAGGCCGAGGAGAAGAAGC	57.8	20	118
	R	GCACGACTCGTAGCATCACA	57.6	20	
<i>Actin (PgAct)</i>	F	CTCAGTGGAGGATCTACTAT	59.4	20	108
	R	GGTGGTGCAATCACTTTAAC	63.2	20	

Where F, indicates the forward primer and R, indicates the reverse primer.

primer was used as an endogenous control to normalize the number of transcripts accumulated for the PgHSP16.97 gene. The least expression value obtained from the control of the whole seedling in the sensitive line was considered as unity for relative quantification, and the relative fold changes in expression for the remaining lines and plant tissues were estimated.

Homology Modeling and Structure Analysis

A three-dimensional (3D) structure of PgHSP16.97 deduced by Modeller v9.11 (Sali and Blundell, 1993) was subjected to the backbone conformation evaluation by Ramachandran plot using Procheck (Laskowski et al., 1993). The final model and the template were subjected to superimposition for a structural comparison using the STRAP interface (<http://www.bioinformatics.org/strap/>).

Genome-Wide *in silico* Database Mining for the Identification of HSP20 Proteins

To identify various other HSP20 genes present in *P. glaucum*, a total of 38,579 proteome sequences of *P. glaucum* were retrieved from the GIGA database (<http://gigadb.org/dataset/100192>). A Hidden Markow Model (HMM) profile of HSP20 (PF00011) downloaded from the Pfam v33.1 database (<https://pfam.xfam.org/family/PF00011>) was queried against the phytozome v.12.1 database of (<https://phytozome.jgi.doe.gov/pz/portal.html>) *Arabidopsis*, rice, maize, sorghum, and foxtail millet. All the resulted sequences which have a hit with expectation (E) <1.0 were retrieved in the construction of reference sequences (Supplementary Table 1). The reference sequences were aligned with default parameters by using the MUSCLE, and an HMM profile was built based on the obtained consensus of the aligned sequences using HMMER tool v3.2. By using the newly generated HMM profile, the proteome database of *P. glaucum* was scanned for probable homologous sequences in *P. glaucum* by using the HMMER search. Pfam and SMART (<http://smart.embl-heidelberg.de/>) were used to further confirm the conserved HSP20 domain. Finally, 28 HSP20s were identified after removing the redundant sequence without the conserved alpha-crystalline domain (ACD) domain of HSP20 and with the molecular weight being outside the range of 11–35 kDa. With the help of the ProtParam online tools (<https://web.expasy.org/protparam/>), the physiochemical properties of HSP20 proteins were predicted.

Chromosomal Location and Phylogenetic Analysis of HSP20

Based on the ascending order of chromosomal position from one end of the chromosome to the other end, PgHSP20 was annotated as PgHSP01 to PgHSP28 and the physical position was constructed by using MapInspect v1.0 (<https://mapinspect.software.informer.com/1.0/>). The representative HSP20 protein sequences from *Arabidopsis*, rice, tomato, and apple together with the PgHSP20 genes predicted from pearl millet were selected for the construction of a phylogenetic tree. A multiple sequence alignment of HSP20 proteins was conducted by using MUSCLE. The protein sequences were imported to MEGA X to construct a phylogenetic tree by using a maximum likelihood (MLE) method.

The parameters utilized are bootstrapped for 1,000-bootstrap replications; substitution model: Jones–Taylor–Thornton (JTT) model with rates among the sites-gamma distributed (G) and partial deletion for gap/missing data treatment.

Digital Expression Analysis of the Identified HSP 20 Genes

To understand the expression pattern of the identified PgHSP20 genes, the digital expression profiling of 28 identified genes was performed by using the Illumina-based whole transcriptome pearl millet RNA-Seq data set available in the NCBI database as the Sequence Read Archive (SRA) database under the accession number SRP151237. The Custom R script based on “annHeatmap” function in R package-Heatplus was utilized to decipher the expression profile of genes under heat stress with the thermotolerant genotype MS 841B and thermo-susceptible genotype PPMI 69 exposed to heat stress for 30 min and 6 h, respectively, at 42°C (Maibam et al., 2020).

Statistical Analysis

The data from the physiological and biochemical parameters were initially analyzed to understand the extent of variability of the traits among the genotypes and the treatments using the ANOVA. The Duncan multiple range test (DMRT) was used for the separation of means within treatments and genotypes at a level of significance, $p \leq 0.05$. One-way ANOVA was carried out to analyze the relative expression levels for various treatments in each genotype. The differences within treatments and genotypes were estimated by using DMRT ($p \leq 0.05$). To compare the relative expression levels of PgHSP16.97 in the lines within/between the treatments for different plant parts, a *t*-test was used. All statistical analyses and assessment of the level of significance were carried out using the R software (v.3.6.2).

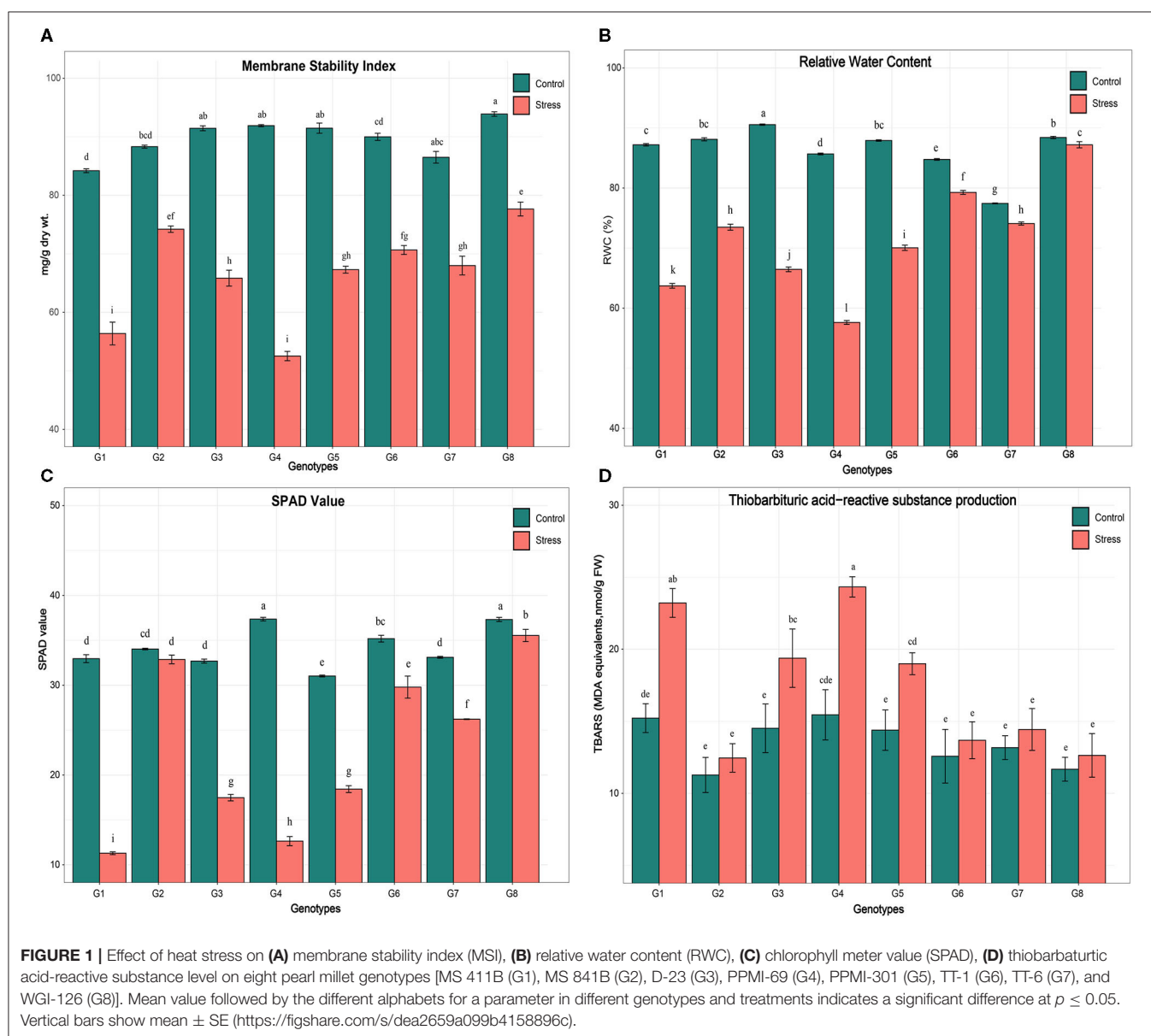
RESULTS

Effects of Heat Stress on Various Physiological and Biochemical Traits of Pearl Millet Genotypes

The ANOVA indicated a high and significant level of genetic variation ($p < 0.001$) among the genotypes for all physiological and biochemical traits under control and heat stress conditions (Table 4). The genotype PPMI-69 (G4) was found to be inferior in all heat stress-related parameters whereas WGI-126 (G8) and MS 841B (G2) showed more endurance to heat stress. Moreover, when heat stress was induced, traits such as MSI, RWC, SPAD values and TBARS levels varied widely. Under heat stress, MSI, RWC, and SPAD readings decreased on average for all genotypes, whereas the TBARS level increased. Furthermore, the treatment–genotype interaction was found to be highly significant ($p < 0.001$), indicating that the genotypes responded to heat stress differently. In pearl millet seedlings, heat stress caused statistically significant ($p < 0.05$) changes in the physiological and biochemical parameters (Figures 1A–D). Although there was a slight variation in the studied traits among the pearl millet genotypes under control conditions, they behaved very

TABLE 4 | The mean square values of the treatments (T), genotypes (G), and the treatment–genotype interaction (T × G) along with their errors and significance indicated by the value of *p*.

Source of variation	Df	MSI		RWC		SPAD		TBARS	
		MSS	Pr(>F)	MSS	Pr(>F)	MSS	Pr(>F)	MSS	Pr(>F)
Treatment (T)	1	6122.45	0.000	2617.90	0.000	1498.96	0.000	178.59	0.000
Genotype (G)	7	282.00	0.000	143.17	0.000	157.11	0.000	59.58	0.000
Replication within Treatment (R)	4	5.55	0.443	0.38	0.254	0.62	0.417	4.30	0.544
T × G	7	97.97	0.000	160.40	0.000	118.35	0.000	15.81	0.021
Residuals	28	5.65		0.27		0.61		5.47	



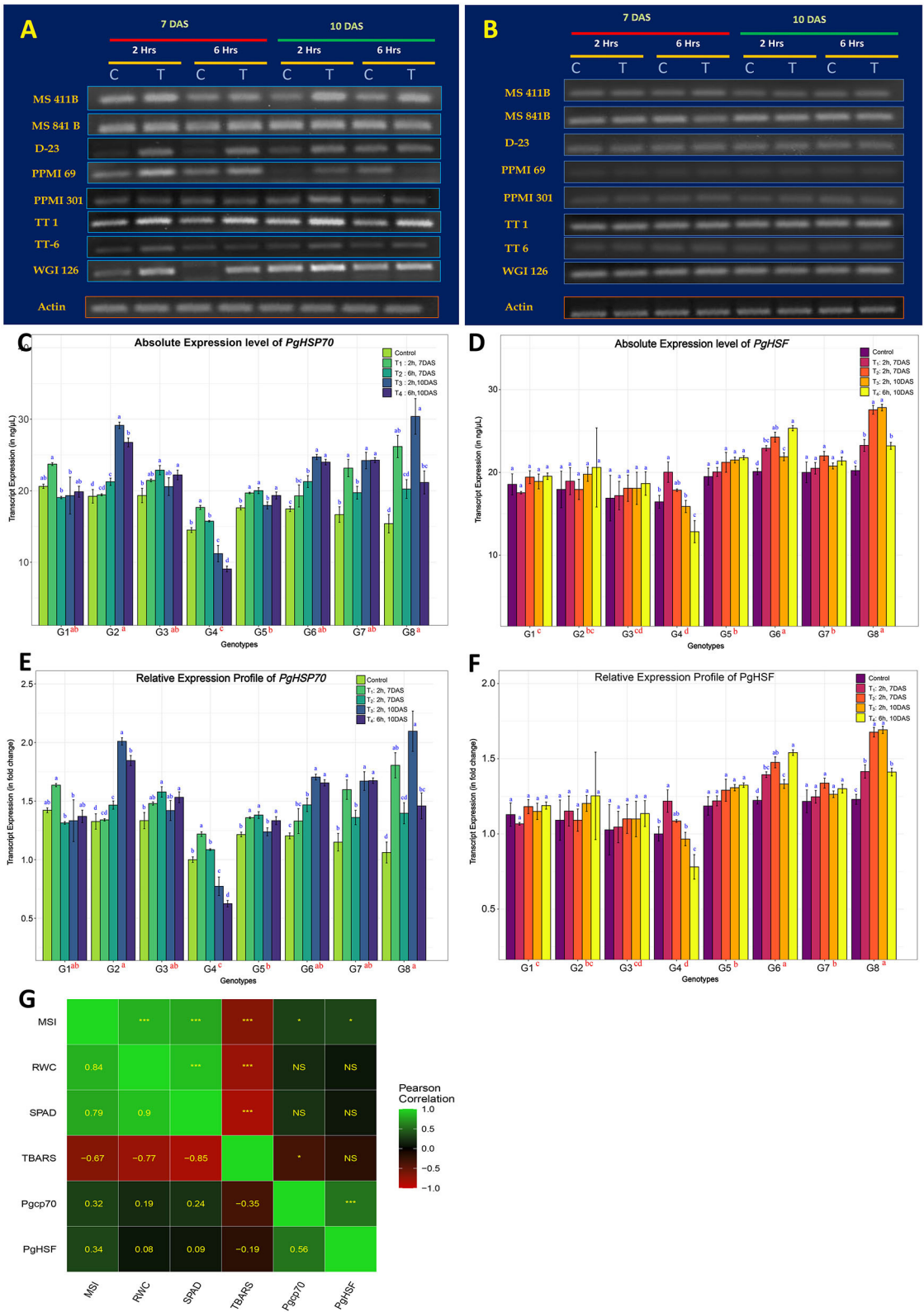


FIGURE 2 | Heat-stress induced gene expression profile of eight pearl millet genotypes for two heat-responsive marker genes. **(A,B)** Reverse transcription PCR (RT-PCR) gel profiles of *Pgcp70* and *PgHSP* genes under regular plant growth (C) and those exposed to heat stress at 42°C (T) during different time periods of treatment indicated above in the seedling stage. The presented results are the representative of at least three independent biological replicates. **(C,D)** Comparative
(Continued)

FIGURE 2 | transcript expression profiling of *Pgcp70* and *PgHsf* for heat tolerance (mean \pm SE) among eight pearl millet genotypes (considering absolute values) after normalization with that of actin gene (internal control). **(E,F)** The relative expression level of *Pgcp70* and *PgHsf* genes (mean \pm SE) among eight genotypes under heat stress. For relative quantification, the least expression value from the whole seedling control in the sensitive genotype (PPMI-69) was considered as unity and the relative fold changes in expression were calculated for the remaining genotype and the treatments. **(G)** Association of marker gene expression values averaged for different heat treatments with physiological and biochemical parameters among different genotypes. Scale on the side of the figure indicates the magnitude and direction of correlations. Shades of green from darker to lighter indicate the strength of a positive correlation between pairs of traits. Shades of red from darker to lighter indicate the strength of a negative correlation between pairs of traits. Darker to black color indicates a very weak or no correlation between a pair of traits. Pearson's correlation coefficients (lower diagonal) and significance (upper diagonal) among different traits were measured in this study. '****' is $p \leq 0.001$, '***' is $p \leq 0.01$, '**' is $p \leq 0.05$, and 'NS' indicates non-significance associationship. Mean sharing the same alphabets for a parameter in different genotypes indicates that there is no significant difference at a significance level of $p \leq 0.05$ (<https://figshare.com/s/0c22f1e7599392059207>).

TABLE 5 | Comparison of the relative gene expression level of marker genes at seedling between different pearl millet genotypes under heat stress treatment (with level significance at $p \leq 0.05$).

S.N	Genotypes	Av.GE.Pgcp70	% change in expression	Stat. Sig*	Rank	Av.GE.PgHsf	% Change	Stat. Sig*	Rank	Cumulative rank
1	MS 411B	1.413	41.3	bc	4	1.147	14.7	c	4	8
2	MS 841B	1.665	66.5	a	1	1.175	17.5	bc	3	4
3	D-23	1.502	50.2	abc	3	1.095	9.5	cd	5	8
4	PPMI-69	0.925	-7.5	d	6	1.013	1.3	d	6	12
5	PPMI-301	1.327	32.7	c	5	1.285	28.5	b	2	7
6	TT-1	1.539	53.9	ab	2	1.436	43.6	a	1	3
7	TT-6	1.576	57.6	ab	2	1.287	28.7	b	2	4
8	WGI-126	1.689	68.9	a	1	1.549	54.9	a	1	2

*Indicates genotypes sharing the same letters are not significant; Av.GE means average gene expression level.

differently when heat stress was imposed. Upon the imposition of heat stress, MSI reduced significantly among all genotypes. However, the genotype G4 (PPMI-69) experienced a major fold reduction in MSI (42%) in comparison with the corresponding control, followed by G1 (MS 411B) and G3 (D-23). On the other hand, G8 (WGI-126) had the least reduction in MSI (17%), followed by MS 841B, upon heat stress.

A similar trend was observed with a significant decrease in RWC in all eight genotypes after imposing heat stress. Genotype G8 (WGI-126) had the least significant reduction in RWC followed by G6 (TT-1) and G2 (MS 841B). Heat stress on seedlings resulted in a significant reduction in chlorophyll content among the genotypes, as indicated by SPAD measurements. Differences were found to be significant ($p \leq 0.05$) in genotypes G4 (PPMI-69) and G1 (MS 411B), under heat stress, with a reduction of 66 and 65%, respectively. On the other hand, G8 (WGI-126) and G6 (TT-1) had the least influence on SPAD reading under stress. A similar trend was followed by the production of TBARS upon oxidative stress due to heat stress in which G4 had the maximum oxidative damage and G8 with the least.

Expression Pattern of Heat-Responsive Marker Genes Among Genotypes Under Different Regimes of Heat Stress

The transcript expression patterns of the two well-studied pearl millet genes, HSP70 and HSF, in response to heat stress were compared to identify the best thermotolerant

and thermosusceptible lines among the selected pearl millet genotypes in the transcript expression level. Responsiveness of genotype to stress in the seedling stage was also established. Hence, young seedlings of 7 days old and 10 days old were chosen, and heat stress of 42°C was imposed for 2 and 6 h, respectively. Comparisons were made within and between the genotypes with and without stress treatment. It was found that both gene transcripts were found to be differentially expressed, in all genotypes irrespective of the treatments and even under controlled conditions (**Figures 2A–F**). *Pgcp70* got expressed under regular growth, the expression level of *Pgcp70* was increased very significantly ($p < 0.05$) upon heat stress. The genotypes showed significant variability for transcript accumulation upon heat stress in which the level was upregulated by 68.9, 66.5, 57.6, 53.9, and 50.2%, respectively, in thermotolerant lines, such as WGI-126 (G8), MS 841B (G2), TT-6 (G7), TT-1 (G6), and D-23 (G3), whereas, in susceptible genotype PPMI-69 (G4), the gene was downregulated by 7.5% to susceptible control (**Table 5**). In MS 411B and PPMI-301, the percent change in expression was comparatively low. The genotype-wise high-temperature tolerance expression pattern of the stress-inducible gene *Pgcp70* can be categorized as WGI-126 \geq MS 841B $>$ TT-1 \geq TT-6 $>$ D-23 $>$ MS 411B $>$ PPMI-301 $>$ PPMI-69 based on the average relative expression level of transcript across the treatments. The expression profiling of *Pgcp70* suggested that HSP 70 was highly induced at an early stage of heat exposure (for 2 h), the transcript level of which was slightly increased as heat stress progressed for a long duration of 6 h on 7-day-old seedlings. Even though we observed a constant induction of HSP 70 during heat stress on

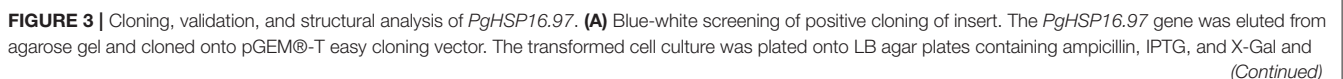


FIGURE 3 | incubated overnight at 37°C. White colonies represent positive cloning. **(B)** Restriction enzyme digestion of the recombinant pGEM®-T easy plasmid with *PgHSP16.97* gene *M* = 1 kb ladder; lane 1 (NR) = non-restricted plasmid; lane 2 (R) = EcoR-I restricted plasmid confirming the presence of insert fragment of size 560 bp. **(C)** The complementary DNA (cDNA) nucleotide sequences, wherein the coding region (letters in red bold font), 5', 3' untranslated region (UTR) regions (letters in black regular fonts). Translation start site and termination codon are underlined, polyadenylation signals (repeated "AT"-rich DNA nucleotide sequences). Various functional domains in the sequence have been significantly marked, such as variable N-terminal domain (light blue font) with hydrophobic groups represented by dark blue font, alpha crystalline domain (ACD; yellow font), and c terminal extension (green font) in which glutamic acid residues and "I X I" residues are represented (in bold green font). **(D)** Schematic representation of *PgHSP16.97* protein structure, including three motifs: the N-Terminal (Blue box), ACD (orange box), and the c-terminal extension (green box), with SXXFD motifs, Conserved regions (CR-I and CR-II) and IXI motifs. **(E)** Alignment of class-I cereal HSP17 amino acid sequences-*PgHsp16.97* (this study), *PgHSP17* (X94191.1), *OshSP 16.9* (AAB39856.1), *TaHSP16.9* (P12810.1), *ZmHsp16.9* (ACG24656.1), *SbHSP16.9* (XP_002457411.1), and *HvHSP16.9* (ADW78607.1) in which SC= start of carboxyl-terminal domain in each alignment, which is the most conserved region of the alpha-crystallin/small heat shock protein (HSP) family; * = conserved residue, : = conserved residue with strongly similar property, = conserved residue of weakly similar property (ClustalW; www.ebi.ac.uk). **(F)** Phylogenetic tree constructed using the ClustalW software showing *PgHSP16.97* and its inferred evolutionary relationships with HSP 17 genes identified from different crops/cereals (<https://figshare.com/s/ef60263530940ec529f4>).

10-day-old seedling, the result suggested that heat stress in the early phase (for 2 h) led to the upregulation of *Pgcp70* transcript in all genotypes though its level is reduced with continuous exposure to heat stress (6 h) over a period. In addition, the transcript accumulation showed a slight increase from 7- to 10-day-old seedlings, indicating that the plant tends to increase its tolerance to high temperature with growth, development, and continued exposure.

In contrast to *Pgcp70*, only thermotolerant genotypes, such as G8 (WGI-126) and G6 (TT-1), and thermosusceptible genotypes, such as G4, were found to have a significant change in the *PgHSF* transcript expression, as indicated by the significance level ($p < 0.05$). The susceptible genotype, on the other hand, only showed a 1.3% increase in gene expression (Table 5). The transcript expression was found to increase upon extending the heat stress in most of the genotypes except PPMI-69. Genotypes can be categorized based on the *PgHSF* transcript abundance during the expression profiling as $WGI-126 \geq TT-1 > TT-6 \geq PPMI-301 \geq MS\ 841B \geq MS\ 411B \geq D-23 > PPMI-69$. Based on the cumulative ranks, WGI-126 and TT-1 were considered as the best thermo-tolerant lines and PPMI-69 was considered as the most thermosusceptible line.

Moreover, the association of various physiological and biochemical parameters with the expression of heat stress marker genes holds a low-to-moderate significant relation ($r > 0.3$, $p < 0.05$) with MSI. TBARS had shown a significant low-to-moderate correlation with the gene expression *Pgcp70*. The physiological parameters, such as MSI, RWC, and SPAD values, have a high positive correlation with each other and show a strong negative correlation with the TBARS level upon heat stress. The expression of two markers upon heat stress was related by $r = 0.56$ (Figure 2G).

In general, considering the physiological and biochemical traits and gene expression of the heat-responsive marker gene, PPMI-69 (G4) was identified as the most thermosusceptible and WGI-126 (G8) as the most thermotolerant genotype among the eight genotypes undertaken in this study.

Isolation and Cloning of Full-Length cDNA of *PgHSP 17.0* From Pearl Millet

Full-length HSP20 ORF of 560 base pairs (bp) and 152 amino acids with an apparent molecular weight of 16.97 kDa and an estimated isoelectric point of 5.79 was isolated and

cloned from the pearl millet heat-tolerant genotype WGI-126, named as *PgHSP16.97*, and deposited with the accession number JQ627835.1 to the NCBI nucleotide database. The homology search that was done by using the deduced amino acid sequence of *PgHSP16.97* against the translated non-redundant nucleotide database suggested that the *PgHSP16.97* was related to the other eukaryotic sHSPs and showed an overall 100–88% sequence identity with sHSPs of *Cenchrus americanus* (Acc no. CAA63901.1), *Z. mays* (Acc no. NP_001150783.1), *Setaria italica* (Acc no. XP_004968025.1), *Saccharum* hybrid cv. ROC22 (Acc no. AFK73383.1). The presence of ACD found in alpha-crystalline-type sHSPs and a similar domain found in p23 (a co-chaperone for HSP90) and in other p23-like proteins confirmed that the isolated sequence belongs to the sHSP gene family (Figure 3).

The ProtComp (<http://linux1.softberry.com/cgi-bin/programs/proloc/protcomp.pl>) analysis produced an integral subcellular localization prediction score of 9.9 for cytoplasmic localization, which indicated that *PgHSP16.97* belongs to class I sHSP. It also carried a nuclear localization sequence. The *PgHSP 16.97* (Figure 3D) monomer contains a variable N-terminal domain (aa, 1–46), the conserved HSP20 or ACD (aa, 47–134), and a less variable C-terminal extension (aa, 135–152). The organelle-localized sHSPs have the necessary transit, targeting, or signal located on N-terminal of protein, which was absent in the sequence-indicating cytoplasmic localization. There is considerable diversity in HSP 17.0 evolution in cereals, with a few conserved motifs and regions, such as IXI motifs, consensus region I (CR-I) (P-X₁₄-GVL), CR-II (P-X₁₅-V-L), R residue at position 114 in C-terminal, and SXXFD motif at N-terminal, which were the signature regions of cytosolic sHSP. Only a few highly conserved domains at N-terminus (21/46) were observed during sequence alignment among cereal class I sHSP. The *PgHSP 16.97* shows close similarity to *PgHSP 17.0* (X94191.1) and is evolutionarily very close to *Zea HSP 16.9* proteins and is more divergent to that of sorghum.

Expression Pattern of *PgHSP16.97*

RNA from six different tissues of the contrasting genotypes, viz., PPMI-69 (thermosusceptible), TT-1, and WGI-126 (thermotolerant genotypes), were collected to analyze and understand the expression pattern of the cloned gene *PgHSP16.97* under different developmental stages. Figure 4A shows the intensity of

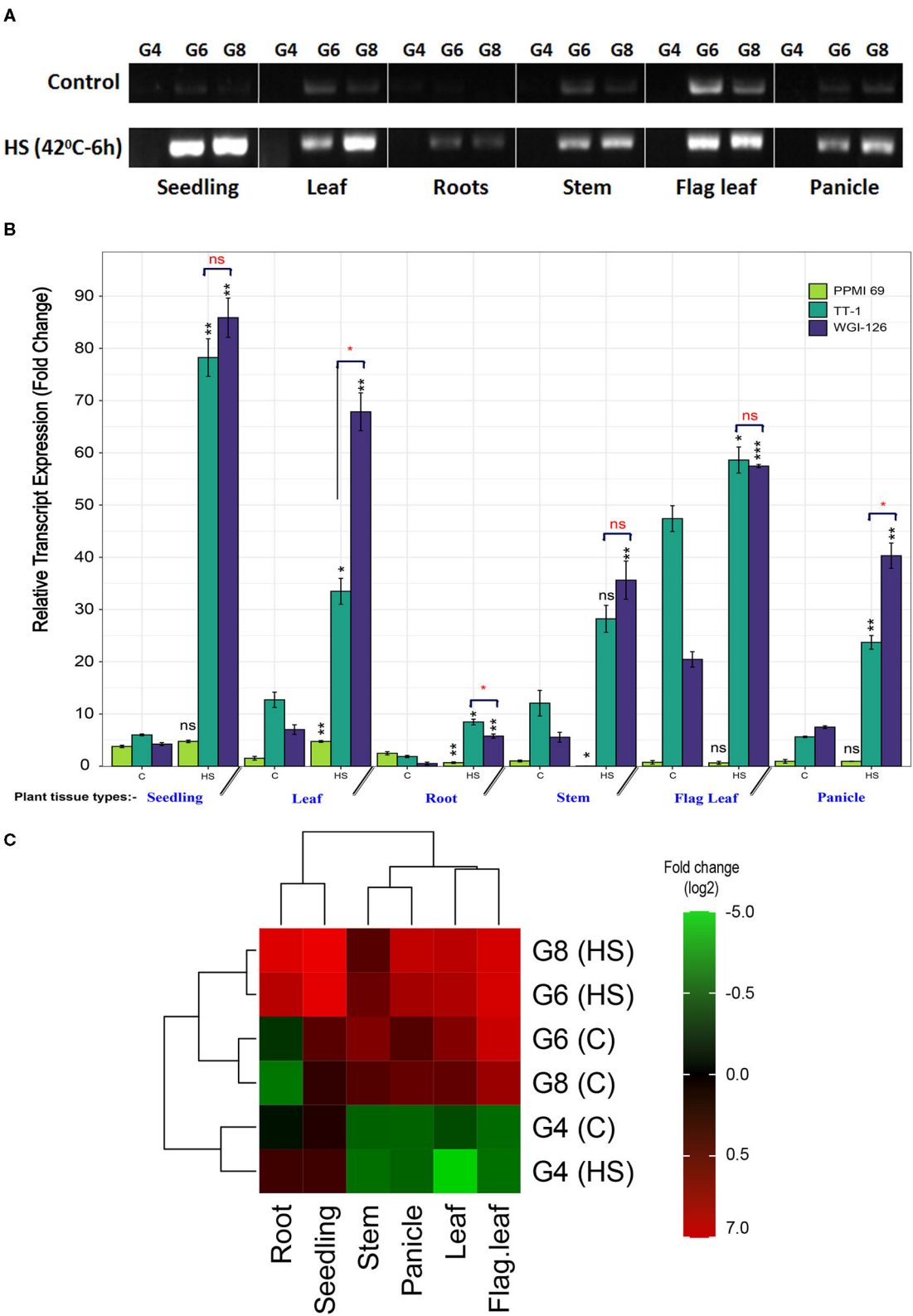


FIGURE 4 | (A) Semiquantitative RT-PCR profile of *PgHSP16.97* transcripts after heat stress (HS) (42°C for 6 h) on different plant tissues of the plant in three contrasting genotypes. The presented results are representative of at least three independent biological replicates. [PPMI-69 (G4), TT-1 (G6), and WGI-126 (G8)].
(Continued)

FIGURE 4 | (B) Relative transcript expression profiling of *PgHSP16.97* for heat tolerance (mean \pm SE) among three contrasting genotypes obtained by quantitative real-time PCR (qRT-PCR). For relative quantification, the least expression value from a control in the sensitive cultivar was taken as unity and the relative fold changes in expression were calculated for the remaining genotypes and the treatments. The paired *t*-test was done to compare the genotypic means between the control and heat-stressed sample to determine the significance of upregulation/downregulation of gene (indicated in black font). Also, the difference in mean among the tolerant genotypes was done to determine a significant difference in the transcript accumulation among the thermotolerant genotypes (red fonts). Significance codes: *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, ns, not significant. **(C)** Heat map of *PgHSP16.97* has been generated based on the log₂-transformed fold-change values in the heat-stressed sample (HS) when compared with its control sample (C). The column of the heat map represents the plant tissue from where the samples were taken, and rows represent a combination of genotypes and treatments. The relative values for each point are depicted by color intensity, with green indicating the downregulation and red indicating the upregulation of *PgHSP16.97* gene. The color scale for log₂-transformed fold-change values is shown in the right side (<https://figshare.com/s/4fbc26be5c90b74f521d>).

TABLE 6 | Comparison of the relative gene expression level of *PgHSP16.97* at different tissues among the contrasting pearl millet genotypes under heat stress treatment.

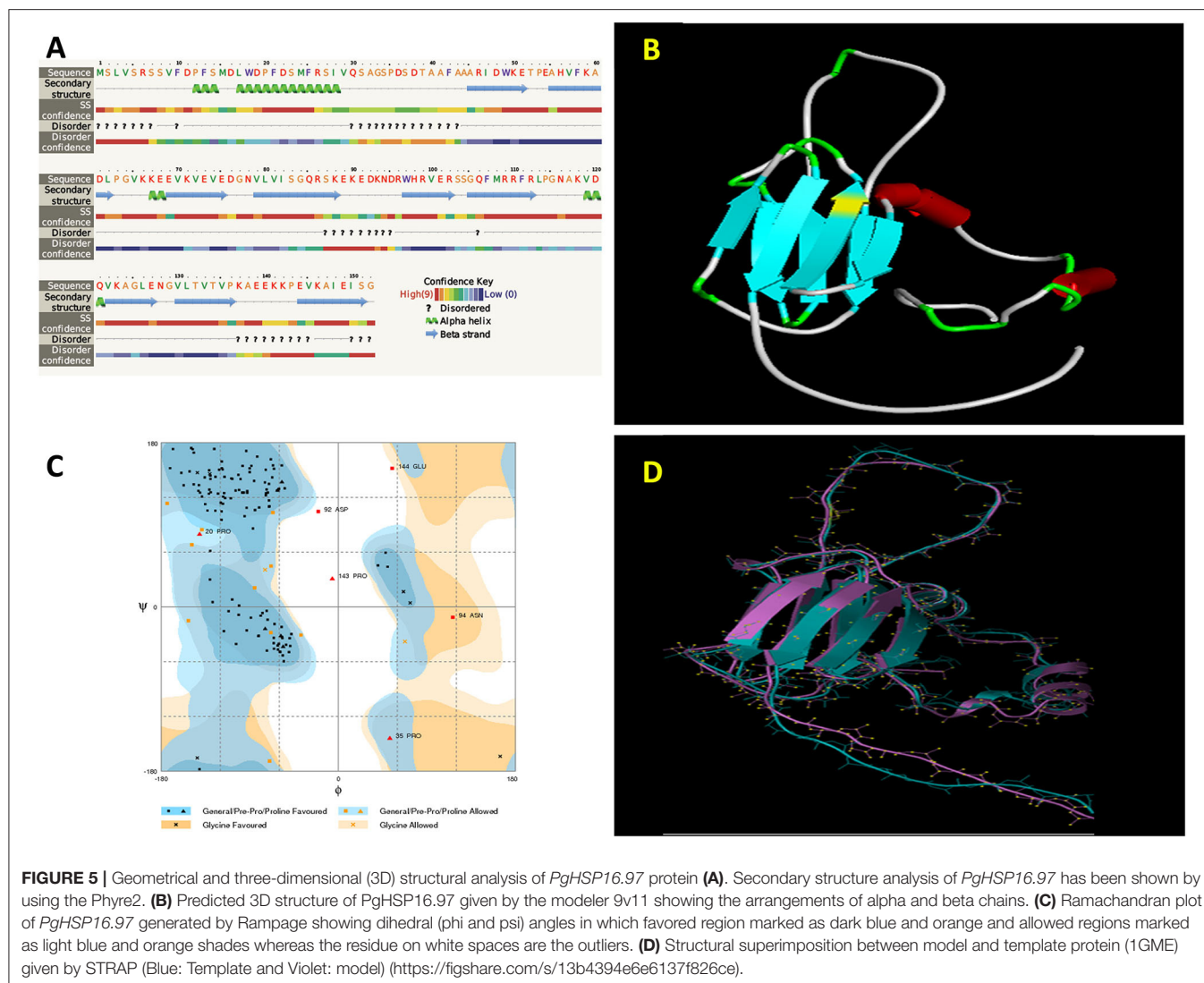
Samples	Genotype	Fold change	Gene expression	Tolerant vs. Suceptible		B/w Tolerant	
				t-stat	p-value	t-stat	p-value
Seedling	PPMI 69	0.26	Upregulation	−3.84	0.062	−1.16	0.366
	TT-1	12.05	Upregulation	−21.15	0.002		
	WGI-126	19.31	Upregulation	−20.33	0.002		
Leaf	PPMI 69	2.11	Upregulation	−12.02	0.007	−5.86	0.028
	TT-1	1.64	Upregulation	−6.96	0.020		
	WGI-126	8.67	Upregulation	−16.24	0.004		
Root	PPMI 69	−0.73	Down regulation	11.32	0.008	5.11	0.036
	TT-1	3.57	Upregulation	−9.13	0.012		
	WGI-126	10.37	Upregulation	−9.73	0.010		
Stem	PPMI 69	−0.96	Down regulation	5.10	0.036	−1.80	0.214
	TT-1	1.34	Upregulation	−3.21	0.085		
	WGI-126	5.43	Upregulation	−11.03	0.008		
Flag leaf	PPMI 69	−0.13	Down regulation	0.16	0.889	0.53	0.652
	TT-1	0.24	Upregulation	−4.14	0.050		
	WGI-126	1.81	Upregulation	−25.92	0.001		
Panicle	PPMI 69	0.00	Constant	0.00	0.999	−4.76	0.041
	TT-1	3.23	Upregulation	−12.49	0.006		
	WGI-126	4.39	Upregulation	−14.05	0.005		

PCR products as measured by densitometry. Upon heat stress, the relative expression level of *PgHSP16.97* increased significantly ($p < 0.05$) in different tissues of genotypes except in seedlings, flag leaf, and panicle of susceptible genotype PPMI-69. It had the least expression of *PgHSP16.97* in different tissues at different stages during qRT-PCR analysis (Figures 4B,C). Among the thermotolerant genotypes, WGI-126 had the highest expression at the seedling stage and was 19 times higher than the expression in the corresponding control. In addition, WGI-126 showed a significant difference in the expression concerning another thermotolerant genotype TT-1, in the case of the transcript expression at leaf, root, and panicle under heat stress. In general, the transcript expression of the *PgHSP16.97* was higher in seedlings (10.5 times), followed by roots (4.4 times), leaves (4.1 times), and least in flag leaves (0.6 times) (Table 6).

Homology Modeling of *PgHSP16.97*

The functionality of a protein largely depends on the structural features such as a motif, domain, and their 3D confirmations. The deduction of its 3D structure from its primary structure has been

considered as one of the major goals to understand the structural dynamics of the protein. Proteins are known to be dynamic entities in cellular solutions. Homology modeling is one of the approaches that are followed for the 3D structure prediction of proteins based on their match with the primary amino sequence of the experimentally solved homologous protein. In our studies, *PgHSP16.97* shared 80% similarity in their primary amino acid sequence levels with the wheat *HSP16.9* protein (PDB no: 1GME). Hence, the crystal structure of the wheat *HSP16.9* protein (PDB no: 1GME) was chosen as a template for building a *PgHSP16.97* model using the program Modeller 9v11 (Sali and Blundell, 1993). The secondary structure of the protein consists of four alpha-helices and nine beta-strands (Figure 5A). The Accelrys Discovery Studio 3.5 Client program was used to depict the *PgHSP16.97* molecular model (Figure 5B). Five models of 3D structures of *PgHsp16.97* were generated at various refinement levels and validated by using the program Procheck. The best model with a Procheck score of −0.09 was selected (Figure 5C). Superimposition of the model with the template and root mean square deviation (RMSD) calculation was carried out using



the program (<http://www.bioinformatics.org/strap/>). The RMSD value of the selected *PgHSP 16.97* model structure is 1.60 Å° regarding the template 1GME. The structural superimposition was done with the STRAP interface and was observed to have a better level of the superimposition of the model onto the template, as shown in **Figure 5D**.

Chromosomal Distribution and Structural Analysis of *PgHSP20* Genes

The present investigation revealed that *PgHSP 16.97* is located on chromosome 6 and their ortholog in the pearl millet genome revealed that 28 genes were evenly distributed across the genome (**Supplementary Table 2**). A variation in gene length of HSP20 varies from 338 bp (*PgHSP03*) to 17.6 kbp (*PgHSP27*). The chromosomal location with its coordinates is shown in **Figure 6**. Chromosomes 2 and 6 have most of the HSP20 genes (7), most of which are located on the telomeric regions, whereas chromosome

7 has the least one. The structural features of the identified genes were elucidated by using the Gene Structure Display Server (GSDS). **Figure 7** shows a varying pattern of total exonic and intronic regions among the 28 *PgHSP20* genes. Among 28 genes, 11 genes (39.28%) possess only a single exon, 10 genes (35.71%) have two exons and a single intron, 14.28% of the genes have three exons and two introns, and the remaining one gene, and *PgHSP02* has multiple exons of five.

Phylogenetic Analysis of *PgHSP20* Gene Family

A phylogenetic tree was constructed based on the amino acid sequences of HSP20 genes (**Figure 8**), 28 from pearl millet (*P. glaucum*), 22 from rice (*O. sativa*), 19 from *Arabidopsis thaliana*, 37 from tomato (*Solanum lycopersicon*), and 41 from apple (**Supplementary Table 3**). A total of 147 HSP20s were divided into 16 subfamilies, 35 cytosol Is (C-Is), 19 C-IIIs, 5 C-IIIs, 3 C-IVs, 6 C-Vs, 8 C-VI, 4 C-VII, 10 mitochondria Is (M-Is), 4

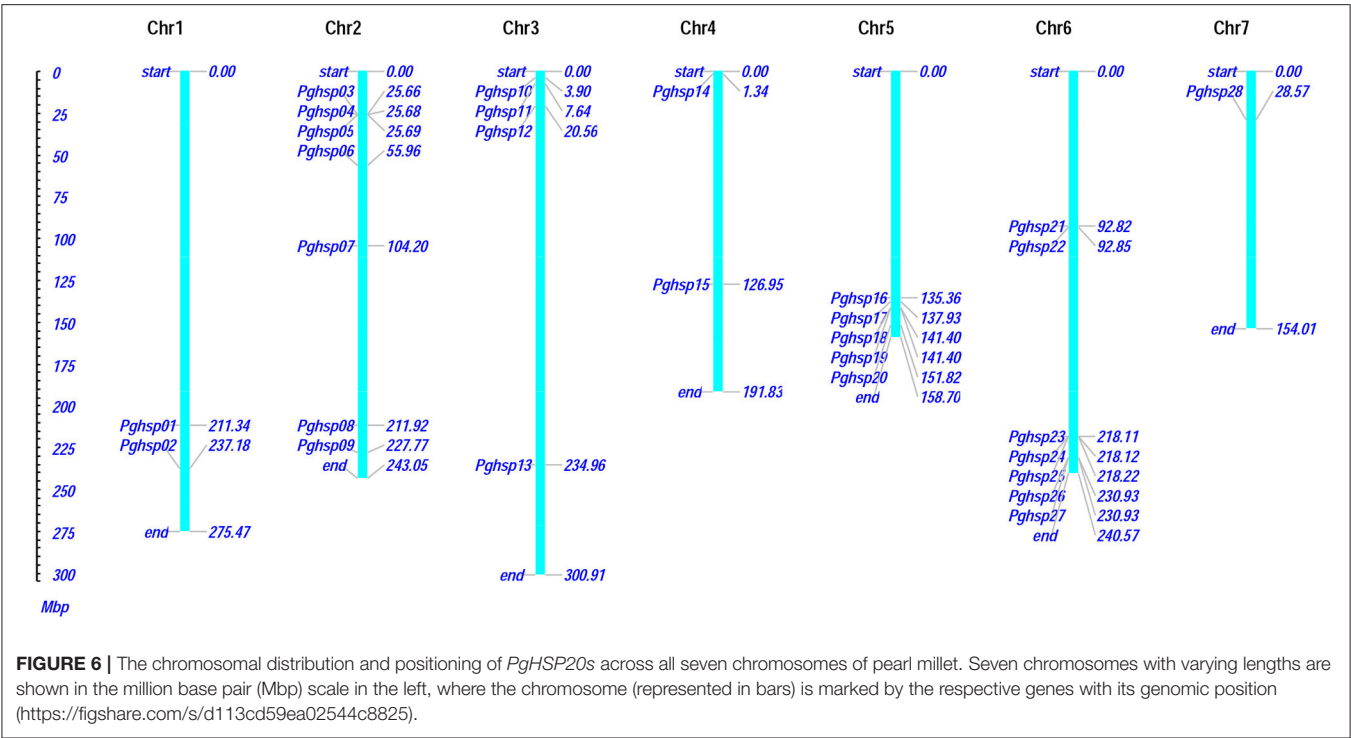


FIGURE 6 | The chromosomal distribution and positioning of *PgHSP20s* across all seven chromosomes of pearl millet. Seven chromosomes with varying lengths are shown in the million base pair (Mbp) scale in the left, where the chromosome (represented in bars) is marked by the respective genes with its genomic position (<https://figshare.com/s/d113cd59ea02544c8825>).

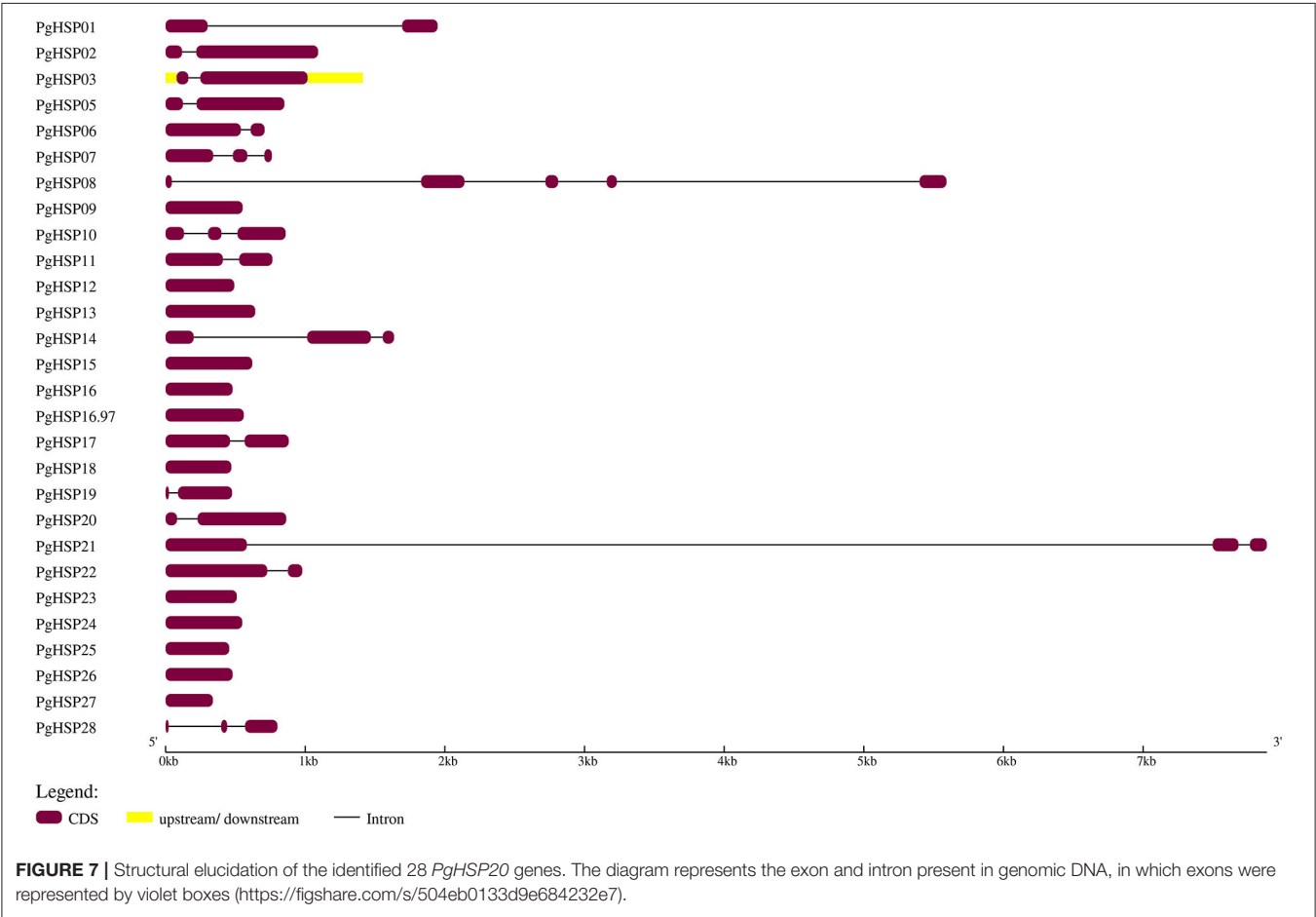
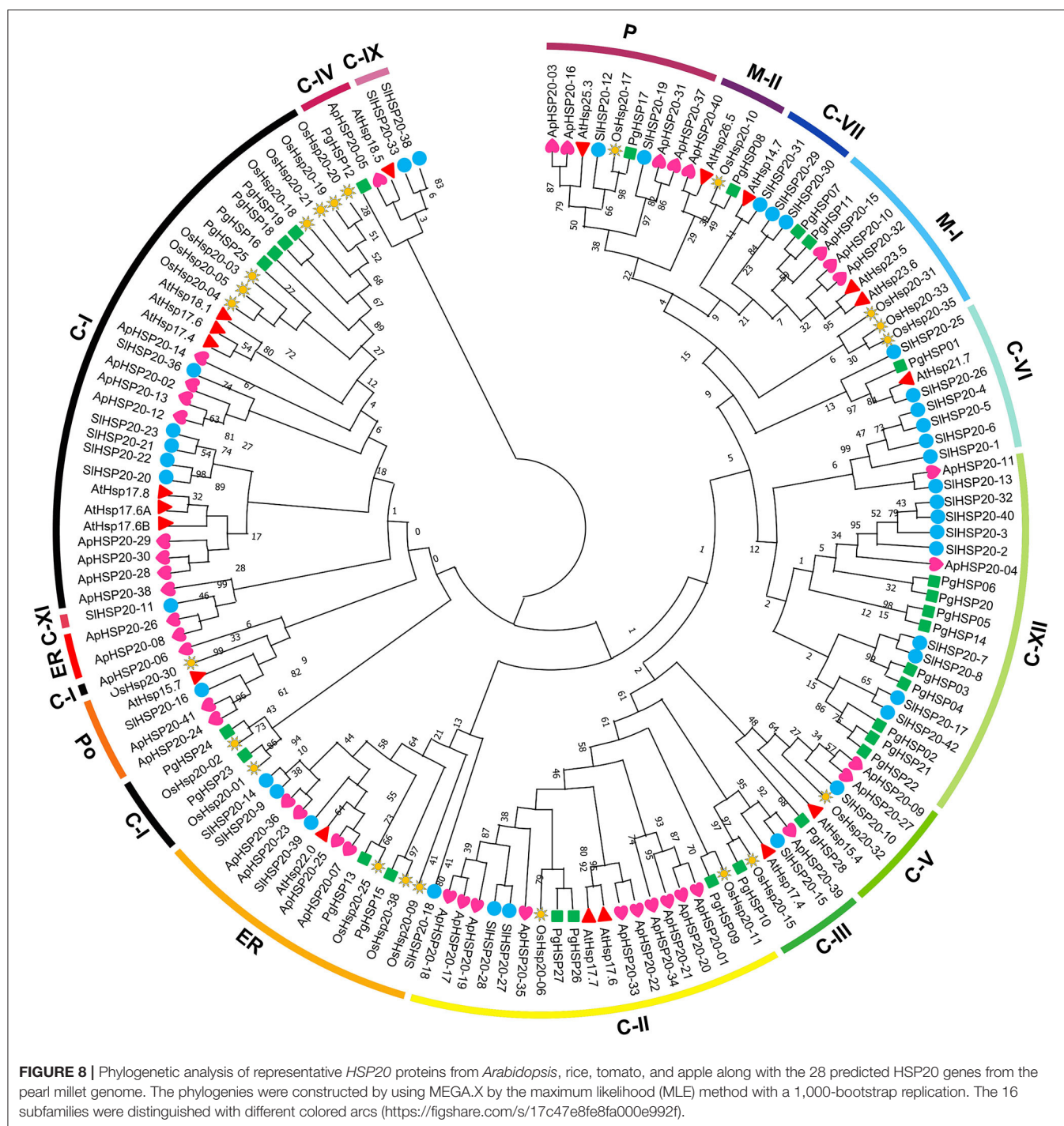


FIGURE 7 | Structural elucidation of the identified 28 *PgHSP20* genes. The diagram represents the exon and intron present in genomic DNA, in which exons were represented by violet boxes (<https://figshare.com/s/504eb0133d9e684232e7>).



M-IIs, 5 peroxisomes (Po), 15 from the endoplasmic reticulum (ER), and 9 from plastids (P) based on the phylogeny and the subcellular localization. The new classes were also divided as follows: 2 proteins in C-IX, 1 in C-XI, and 20 in C-XII. However, the PgHSP20 proteins were distributed into 12 out of 16 subfamilies, in which nine proteins were in C-XII, six in C-I, three in C-II, two each in M-I and ER, one each in C-III, C-IV, C-V, C-VI, M-II, and P based on the phylogeny and the subcellular localization. Out of 28 HSP20s, 22 belonged to the

cytosol, which suggested that the cytosol might be the functional site of the plant HSP20s.

Digital Expression Profile of PgHSP20s in Response to Heat Stress

To further explore the expression profiles of PgHSP20 genes in response to heat stress, whole-genome RNA-Seq libraries, including three independent biological replicates of two contrasting genotypes with the control and heat-treated set,

were downloaded from NCBI (SRP151237) and analyzed. The Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values from the RNA-Seq data were used to estimate the expression levels of 28 pearl millet *HSP20* genes and were further employed in the creation of expression heat map (Figure 9). The expression analysis of the contrasting genotypes of pearl millet under the various degrees of heat stresses has shown a highly uneven expression pattern by most of the *PgHSP20* genes. The heat map also showed that the 28 *HSP20* genes clustered in three groups. Cluster A contains five members (HSP20-16, 18, 24, 25, and 28), which were highly expressed upon heat stress compared with the control

in both tolerant and susceptible genotypes. All 19 members from cluster B were mainly upregulated after 6 h of heat stress in stress-tolerant genotypes alone. The remaining four members (8, 13, 14, and 21) grouped in cluster C were barely upregulated upon heat stress as compared to the control in both contrasting genotypes.

DISCUSSION

Heat shock proteins are a collection of molecular chaperones that have been shown to have a critical role in plant stress reaction as well as growth and development of plants. Small HSPs are the

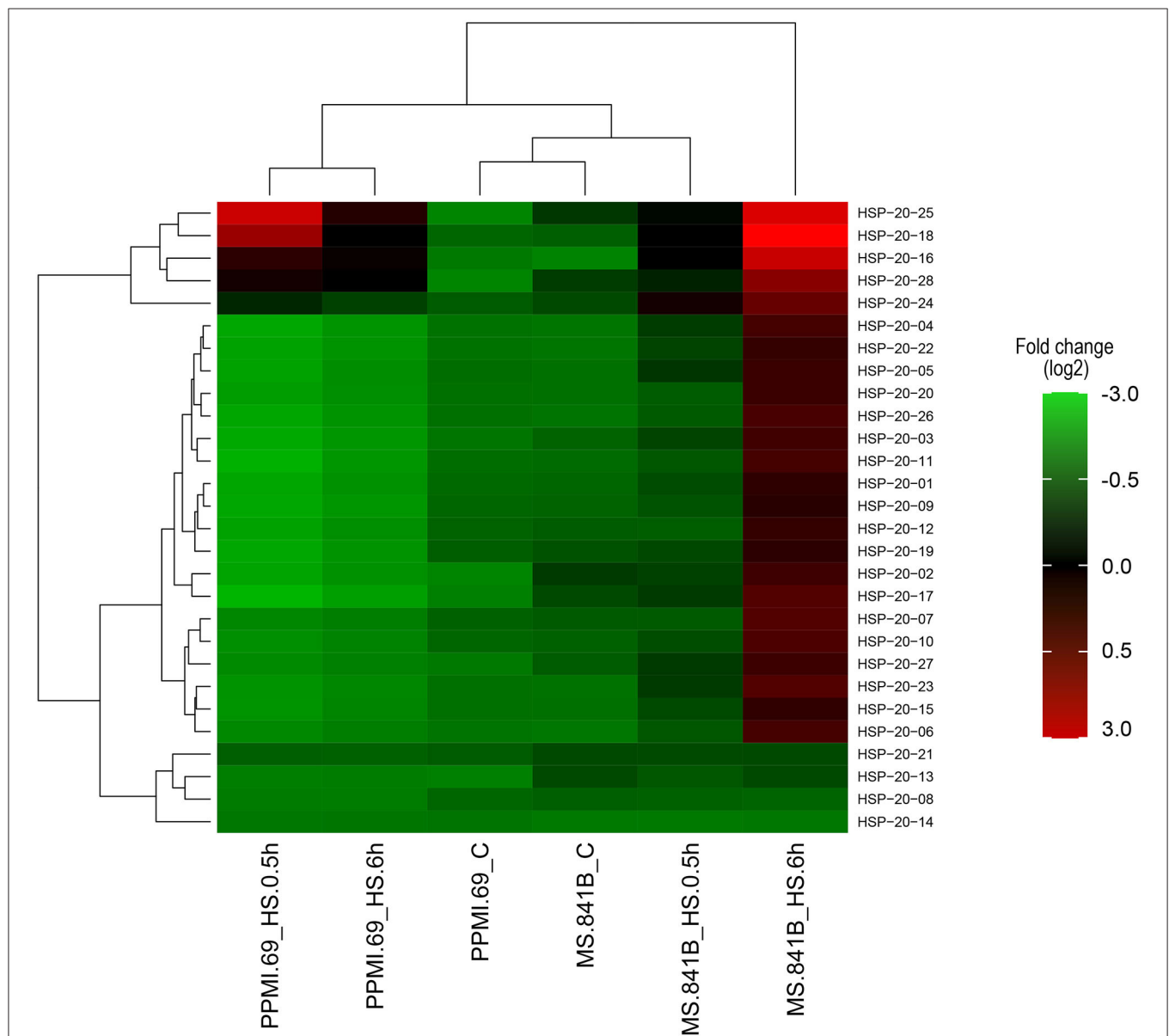


FIGURE 9 | The expression heat map of 28 pearl millet *HSP20* genes under heat stresses based on the RNA-sequencing (RNA-Seq) data. The heat map with clustering was created by using a log₂-based Fragments Per Kilobase of transcript per Million mapped reads (FPKM) value. On the right side of the heat map with a color scale representing relative expression values. C, control plants were maintained at 25°C; HS, heat stress of 42°C for 0.5 and 6 h, respectively, was imposed on plants maintained in the growth chamber (<https://figshare.com/s/7132de92e289274bc8e1>).

largest and diverse group of HSPs among the plants that possess definite roles in various abiotic stress responses (Scharf et al., 2001; Ouyang et al., 2009; Sarkar et al., 2009; Yu et al., 2016; Zhao et al., 2018; He et al., 2019; Ji et al., 2019; Yao et al., 2020), but they are poorly studied in case of pearl millet.

Response of Pearl Millet Genotypes to Heat Stress

The study was initiated with an analysis of some selected physio-biochemical parameters among the eight pearl millet genotypes. Significant differences were observed among the genotypes, treatments, and their interaction for all the studied physio-biochemical parameters. Upon heat stress, the value of MSI, RWC, and SPAD was diminished among the genotypes and prominent in thermosusceptible genotypes, such as PPMI-69, whereas the level of lipid peroxidation increased among the genotypes upon heat stress due to the formation of free radicals and ROS. Thermo-tolerant genotypes, such as WGI-126 and TT-1, showed limited changes to the abovementioned parameters in response to heat stress and maintaining cellular homeostasis. High genetic variability among the pearl millet genotypes using physiological parameters, such as MSI, RWC, and TBARS, for seedling thermotolerance was observed in earlier reports by Mukesh Sankar et al. (2014), James et al. (2015), and Maibam et al. (2020).

Heat stress leads to the induction of various thermotolerant genes, such as HSFs and HSPs, in pearl millet genotypes. Earlier research experience on the responses of HSP70 and HSF in pearl millet seedlings to heat stress in the gene expression level (Maheswar Rao et al., 2010; Reddy et al., 2010; Divya et al., 2019; Sun et al., 2020) suggested that heat stress response could be used as a molecular marker to identify heat-responsive genotypes. In the current experiment, the expression pattern suggested the upregulation of the transcript level of heat-responsive genes, such as *Pgcp70* and *PgHSF*, under high-temperature stress, supporting the earlier findings of the heat-induced HSPs (Howarth. C. J., 1991; Mishra et al., 2007; Sun et al., 2020). It was also noticed that, under a normal growth condition (25°C), there was a slight accumulation of these transcripts as HSP 70 and its master regulator, HSFs, have a critical role in seed germination and development and also could be attributed to the inherent thermotolerance in pearl millet cultivars as shown by the result of the MSI (Maibam et al., 2020). The *Pgcp70* was expressed in small quantity under normal conditions but was enhanced with heat treatment. A large quantity of HSP 70 in the early stage of heat exposure (2 h) indicates a major role in heat stress during the early phases of heat stress. Later on, the transcript level decreases as stress is continued (6 h), which indicates that an immediate shock in the genotype leads to a higher and rapid induction of *Pgcp70*, the level of which diminishes with continuous exposure to heat stress, suggesting constant involvement of HSP in heat shock for a longer duration (Waters et al., 1996). It was also noticed that the *Pgcp70* transcript accumulation showed a slight increase from 7- to 10-day-old seedlings, which indicated that plants tend to increase the temperature tolerance with their development. The expression of the *PgHSF* was at much lower

levels when compared to the transcript level of *Pgcp70*. Even such a low-level HSF transcript could trigger the transcription of *Pgcp70* under heat stress. The transcript expression of the HSF gene in pearl millet showed an utmost rise in the transcript level within 30 min of exposure and gradually came down with time upon heat stress. This indicated its stress regulatory or initiator role. This result was in agreement with the study of Reddy et al. (2010). The expression profile of *PgHSF* was nearly constant for a long period of heat exposure for 2 h and 6 h in comparison with the expression profile of *Pgcp70*. *PgHSF* transcript abundance under high-temperature treatment in the seedling stage was less than *Pgcp70* with the same condition, which suggests an initiator role towards rapid stress response in germinating seedling.

The expression of *Pgcp70* and *PgHSF* was found to be associated with the changes in MSI and TBARS under heat stress. Both parameters had shown a low-to-moderate significant correlation with the gene expression of both marker genes. Some studies suggested that the HSP co-inducer hydroxamic acid derivatives intercalate into the membrane and stabilize their lipid rafts by regulating membrane structure and composition (Balogi et al., 2019). Similarly, Usman et al. (2018) reported the role of HSP70 in membrane stability in chili pepper. Also, the accumulation of ROS upon heat stress leads to the accumulation of thiobarbituric acid derivatives leading to membrane lipid damages, and subsequently the electrolyte leakage (Killi et al., 2020). The extent of differences observed in HSP70 expression, membrane stability, and lipid peroxidation can be used to improve the thermotolerance capacity of the genotypes. The differential expression pattern was observed for these two heat-responsive genes among different genotypes. WGI-126 showed a positive response to heat stress by accumulating more transcripts whereas PPMI-69 exhibited a very low accumulation of transcripts, as indicated by the physio-biochemical parameters. These results suggested that transcript expression profiling can be used for screening thermo-tolerant and susceptible lines in a large pool of genotypes along with the identification of the potential genes responsible for thermotolerance at various stages of crop growth.

Isolation and Characterization of *PgHSP16.97* Under Heat Stress

Among the molecular chaperones, the sHSPs were diverse and were found in both prokaryotes and eukaryotes, which were usually untraceable under a normal physiological situation in plant cells, but they were induced upon stress, leading to plant tolerance to stress, such as drought, poor plant nutrient availability, salinity, and extreme temperatures (Singh et al., 2016). Hence, the diversification and abundance of sHSPs in a plant are thought to reflect the adaptation of a plant to heat stress (Pareek et al., 2021). Among the HSPs, HSP20 forms the first line of defense against stress in the cell during heat stress. They have the ability to bind to the denatured or partially folded proteins to avoid irreversible unfolding or erroneous protein aggregation or to bind to unfolded proteins in an energy-free manner until suitable conditions for renewed cell activity arise, allowing further refolding by HSP70/HSP100 complexes,

earning the nickname “paramedics of the cell” (Hilton et al., 2013). Hence, we have isolated and cloned full-length cDNA encoding for HSP 17.0 from *P. glaucum* to understand the structural signature present in this protein for its role in heat tolerance among genotypes. The analysis of nucleotide and deduced amino acid sequence of the cDNA clone revealed the presence of alpha-crystalline-type sHSPs. A similar domain found in p23 (a co-chaperone for Hsp90) and other p23-like proteins confirmed that the isolated sequence belongs to the sHSP gene family. Alpha-crystalline occurs as large aggregates, comprising two types of related subunits (A and B) that are highly similar to the small HSPs (11–35 kDa), particularly in their C-terminal halves. Alpha-crystalline has chaperone-like properties, such as the ability to protect proteins from precipitation, thereby improving the cellular tolerance to stress. The model structure revealed that the N-terminal arm of the *PgHSP16.97* represents an extensive, intrinsically unstructured domain rich in hydrophobic residues (53%), which will play key roles in protein–protein interactions with denatured proteins and thus critical to substrate interactions (Jaya et al., 2009). Due to its ability to present multiple binding site conformations, *PgHSP16.97* is highly effective at interacting efficiently to protect a wide range of critical cellular proteins. In addition, N-terminal regions are important for stabilizing an oligomer through interlocking subunits by forming two disks intertwined to form pairs of knot-like structures, and the hydrophobic contacts in these knots are buried inside an oligomer (van Montfort et al., 2001). The C-terminal extension is variable in length, and its function in those cellular compartments is enigmatic. As per the sequence information, the C-terminal extension was rich in glutamic acid residues (E-), which were critical for its chaperone activity (Morris et al., 2008). Moreover, there was an Ile-X-Ile residue at the C-terminal extension (β_{10} strand), which has a role in the oligomerization of HSPs (Studer et al., 2002) by interacting with the hydrophobic pockets formed at β_4 and β_8 of ACD strands. Two CRs within C-terminal separated by a region with hydrophilic residues form a signature sequence for the identification of cytosolic plant sHSP. The CR-I consists of residues (P-X₁₄-GVL) that are involved in the multimerization of *PgHSP17* subunits, and the CR-II consists of residues (P-X₁₅-V-L) involved in the solubility of the protein complex. Arginine (R) conserved across the cereal sHSP gene at position 114 is responsible for the stabilization of dimer by the formation of an intermolecular salt bridge with glutamic acid at position 100 (van Montfort et al., 2001).

Moreover, the gene expression pattern in pearl millet suggested its constant involvement in response to heat stress at different developmental stages. Under controlled conditions, the expression level of *PgHSP16.97* was very low, still showing a higher expression in flag leaf, stem, and developing panicles, which indicated the involvement of HSP20 genes in plant reproductive tissue development. Cytosolic class I HSP20s will begin to accumulate on flag leaf and develop seed from mid-maturation throughout the late maturation under normal conditions. A similar observation concerning the HSP20 genes was reported by Poidevin et al. (in press) while conducting riboprofiling studies on developing pollens and in developing

wheat panicles (Wang et al., 2020). Upon heat stress, the expression of the gene was upregulated by up to 19 times in almost all tissues of tolerant genotypes. However, the expression was found to be higher in the seedling stage as compared to other tissues. Young seedlings are prone to dehydration by heat stress than other tissues due to the presence of more meristematic tissues. Several studies have speculated that cytosolic class I HSP20s may function to protect cellular components during seedling desiccation (Kalemba et al., 2019).

Genome-Wide Scan and Expression Profiling of Pearl Millet HSP20 Gene Family in Response to Heat Stress

In the current investigation, we used the pearl millet genome database to identify 28 HSP20 genes and investigated their sequence characteristics. The number of HSP20 genes in pearl millet was higher than that in *Arabidopsis* (19) (Scharf et al., 2001), but slightly lower than that in tomato (37) (Yu et al., 2016), foxtail millet (37) (Singh et al., 2016), and rice (39) (Ouyang et al., 2009), and much lower than that in sorghum (47) (Nagaraju et al., 2020) and bread wheat (163) (Muthusamy et al., 2017). The gene duplications and rearrangements that occurred during evolution are quicker to attribute for this difference (Devos et al., 2000; Varshney et al., 2017). The expansion of the number of gene families in plants is thought to be aided by gene duplication and the same with the evolution of sHSPs (Waters and Vierling, 2020). The pearl millet HSP20 family could be divided into 12 subfamilies according to the phylogenetic analysis, which is consistent with the previous evolutionary classification of HSP20 genes in sorghum and foxtail millet (Singh et al., 2016; Nagaraju et al., 2020). A total of 22 out of 28 members belong to nucleo-cytoplasmic subfamilies, and among these subfamilies, CXII was the largest subfamily with 20 members. As a result, we hypothesized that the cytoplasm, as a site primarily for protein synthesis, could be the primary location for HSP20 interacting with the denatured proteins to prevent them from aggregating and degrading inappropriately (Waters and Vierling, 2020). The gene structure analysis indicated that most pearl millet HSP20 genes (74.99%) have no introns or a single intron, suggesting relatively simple gene structures. The gene structure analysis indicated that nearly 75% of the pearl millet HSP20 genes were either intronless or with a single intron, implying relatively simple gene structures. This is concurrent with the results of the previous finding of the researchers, which showed that the plants were prone to retain more genes with no intron or a short single intron (Ji et al., 2019). According to the RNA-Seq atlas, temporal regulation of the *PgHSP20* gene family was observed under heat stress among the contrasting genotypes. The heat map showed that the *PgHSP20* genes clustered in three groups, in which cluster A contains five members (HSP20-16, 18, 24, 25, and 28), which were highly expressed upon heat stress compared with the control in both tolerable and susceptible genotypes. These genes have no introns except for *PgHSP-28*. Furthermore, we also found that these *PgHSP20* genes were more heat stress-inducible in susceptible genotypes compared to tolerable genotypes, as evidenced by their expression pattern after 30 min of heat stress

induction. The rapid response of sHSPs upon heat stress in susceptible genotypes over tolerant genotypes for a short period of stress was also observed in crops (Chakraborty et al., 2018).

CONCLUSION

HSP20 genes are a large and diverse group of genes involved in plant stress responses, and as a result, they have been studied extensively in a variety of crop plants. However, to the best of our knowledge, there are a handful of reports available in otherwise naturally stress-tolerant model C₄ crops such as pearl millet. An attempt was made in the current study to validate the powerfulness of the expression pattern study as a molecular screening technique in identifying the contrasting genotypes based on the expression of heat-responsive marker genes, such as HSP or HSF, in the seedling stage of pearl millet. Using the transcript expression-based screening, one gene (PgHSP16.97) belonging to the HSP20 family was identified, cloned, and characterized for the first time from pearl millet cv. WGI 126. Later, this sequence information was used in the identification of the genome-wide distribution of HSP20 genes. Isolating and identifying these functional genes will help researchers in gaining an in-depth understanding of the molecular genetic basis of pearl millet stress adaptation and genetic improvement. It can serve as powerful genomic resources for genetic engineering. Also, the *in silico* structure prediction can provide further insights into the molecular regulation of stress tolerance, thereby bridging them together to fight against the unpredicted nature of abiotic stress.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

REFERENCES

- Balogi, Z., Multhoff, G., Jensen, T. K., Lloyd-Evans, E., Yamashita, T., Jäätelä, M., et al. (2019). Hsp70 interactions with membrane lipids regulate cellular functions in health and disease. *Prog. Lipid Res.* 74, 18–30. doi: 10.1016/j.plipres.2019.01.004
- Bharadwaj, N., Barthakur, S., Biswas, A. D., Das, M. K., Kour, M., Ramteke, A., et al. (2019). Transcript expression profiling in two contrasting cultivars and molecular cloning of a SKP-1 like gene, a component of SCF-ubiquitin proteasome system from mungbean *Vigna radiata* L. *Sci. Rep.* 9, 1–17. doi: 10.1038/s41598-019-44034-4
- Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., et al. (2019). MADS-box genes are key components of genetic regulatory networks involved in abiotic stress and plastic developmental responses in plants. *Front. Plant Sci.* 10:853. doi: 10.3389/fpls.2019.00853
- Chakraborty, K., Bishi, S. K., Singh, A. L., Zala, P. V., Mahatma, M. K., Kalariya, K. A., et al. (2018). Rapid induction of small heat shock proteins improves physiological adaptation to high temperature stress in peanut. *J. Agron. Crop Sci.* 204, 285–297. doi: 10.1111/jac.12260
- Choi, S. H., Park, N., Lee, K. Y., Missaoui, A. M., and Lee, G. J. (2019). Novel genes in response to varying water deficit in oil crop *Camelina sativa*. *Euphytica* 215:86. doi: 10.1007/s10681-019-2402-9

AUTHOR CONTRIBUTIONS

SM performed the investigation and wrote the manuscript. CT and SB were involved in the conceptualization, supervision, and editing of the manuscript. SB helped in validation, funding acquisition, and editing of the manuscript. SPS and CB were involved in the editing of manuscript. SLS performed the data analysis and edited the manuscript for final submission. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the DBT funded project (code: BT/PR13181/GBD/27/215/2009) on Molecular cloning and functional characterization of annexin family genes from Pearl millet (*Pennisetum glaucum*) under abiotic stress.

ACKNOWLEDGMENTS

This study was a part of MSc Thesis and SM wishes to thank ICAR-IARI, New Delhi and ICAR-NIPB, New Delhi, India, for providing the facilities for the study and the JRF provided by ICAR, India. We would like to thank all the review participants in bioRxiv and Frontiers in Plant Science platform who have given critical suggestion in improving the manuscript. The pre-print of the article is now available through the following link, Mukesh Sankar et al. (2020) (<https://www.biorxiv.org/content/10.1101/2020.02.26.966184v1.full>).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.659893/full#supplementary-material>

- Devos, K. M., Pittaway, T. S., Reynolds, A., and Gale, M. D. (2000). Comparative mapping reveals a complex relationship between the pearl millet genome and those of foxtail millet and rice. *TAAG* 100, 190–198. doi: 10.1007/s001220050026
- Divya, K., Kavi Kishor, P. B., Maraka, N., Bhatnagar-Mathur, P., Singam, P., Vadez, V., et al. (2019). Genome-wide Identification and Characterization of Hsp70 gene family in Pearl millet (*Pennisetum glaucum*). *Curr. Trends Biotechnol. Pharm.* 13, 102–111. doi: 10.13140/RG.2.2.13043.17449
- Galvez, L. C., Koh, R. B. L., Barbosa, C. F. C., and Aquino, V. M. (2020). Multiplex reverse transcription-polymerase chain reaction for simultaneous detection of banana bract mosaic virus (BBMV) and sugarcane mosaic virus (SCMV) in abaca. *Can. J. Plant Pathol.* 42, 572–583. doi: 10.1080/07060661.2019.1708810
- He, Y. J., Fan, M., Sun, Y. Y., and Li, L. L. (2019). Genome-wide analysis of watermelon HSP20s and their expression profiles and subcellular locations under stresses. *Int. J. Mol. Sci.* 20:12. doi: 10.3390/ijms20010012
- Heath, R. L., and Packer, L. (1968). Photoperoxidation in isolated chloroplasts I. Kinetics and stoichiometry of fatty acid peroxidation. *Arch. Biochem. Biophys.* 125:189. doi: 10.1016/0003-9861(68)90654-1
- Hilton, G. R., Lioe, H., Stengel, F., Baldwin, A. J., and Benesch, J. L. (2013). Small heat-shock proteins: paramedics of the cell. *Top. Curr. Chem.* 328, 69–98. doi: 10.1007/128_2012_324

- Howarth, C. J. (1991). Molecular responses of plants to an increased incidence of heat shock. *Plant Cell Environ.* 14, 831–841. doi: 10.1111/j.1365-3040.1991.tb01446.x
- Indiastat (2020). *Agricultural Production*. New Delhi: Indiastat, Datanet India Pvt. Ltd. Available online at: <http://www.indiastat.com/searchresult.aspx> (accessed February 27, 2020).
- James, D., Tarafdar, A., Biswas, K., Sathyavathi, C. T., Padaria, J. C., and Kumar, P. A. (2015). Development and characterization of a high temperature stress responsive subtractive cDNA library in Pearl Millet (*Pennisetum glaucum* L. R. Br.). *Indian J. Exp. Biol.* 53, 543–550.
- Jaya, N., Garcia, V., and Vierling, E. (2009). Substrate binding site flexibility of the small heat shock protein molecular chaperones. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15604–15609. doi: 10.1073/pnas.0902177106
- Ji, X. R., Yu, Y. H., Ni, P. Y., Zhang, G. H., and Guo, D. L. (2019). Genome-wide identification of small heat shock protein (HSP20) gene family in grape and expression profile during berry development. *BMC Plant Biol.* 19:433. doi: 10.1186/s12870-019-2031-4
- Kalemba, E. M., Bagniewska-Zadworna, A., Suszka, J., and Pukacka, S. (2019). Dehydration sensitivity at the early seedling establishment stages of the European beech (*Fagus sylvatica* L.). *Forests* 10:900. doi: 10.3390/f10100900
- Killi, D., Raschi, A., and Bussotti, F. (2020). Lipid peroxidation and chlorophyll fluorescence of photosystem II performance during drought and heat stress is associated with the antioxidant capacities of C₃ sunflower and C₄ maize varieties. *Int. J. Mol. Sci.* 21:4846. doi: 10.3390/ijms21144846
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1993). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–496. doi: 10.1007/BF00228148
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Ma, P., Li, J., Qi, L., and Dong, X. (2021). The archaeal small heat shock protein hsp17.6 protects proteins from oxidative inactivation. *Int. J. Mol. Sci.* 22:2591. doi: 10.3390/ijms22052591
- Maheswar Rao, J. U., Reddy, P. S., Mishra, R. N., Gupta, D., Sahal, D., Tuteja, N., et al. and Reddy, M. K. (2010). Thermo and pH stable ATP-independent chaperone activity of heat-inducible Hsp70 from *Pennisetum glaucum*. *Plant Signal. Behav.* 5, 110–121. doi: 10.4161/psb.5.2.10547
- Maibam, A., Nigombam, S., Vishwakarma, H., Lone, S. A., Gaikwad, K., and Mithra, S. A. (2020). Transcriptome analysis of two contrasting genotypes of pearl millet to gain insight into heat stress responses. *Res. Square [preprint]* doi: 10.21203/rs.3.rs-23605/v2
- Mishra, R. N., Reddy, P. S., Nair, S., Markandeya, G., Reddy, A. R., and Sopory, S. K. (2007). Isolation and characterization of expressed sequence tags (ESTs) from subtracted cDNA libraries of *Pennisetum glaucum* seedlings. *Plant Mol. Biol.* 64, 713–732. doi: 10.1007/s11103-007-9193-4
- Morris, A. M., Treweek, T. M., Aquilina, J. A., Carver, J. A., and Walker, M. J. (2008). Glutamic acid residues in the C-terminal extension of small heat shock protein 25 are critical for structural and functional integrity. *FEBS J.* 275, 5885–5898. doi: 10.1111/j.1742-4658.2008.06719.x
- Mukesh Sankar, S., Satyavathi, C. T., Barthakur, S., Singh, S. P., Kumar, R., Prabhu, K. V., et al. (2020). Differential modulation of heat inducible genes across diverse genotypes and molecular cloning of a sHSP from Pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *bioRxiv [preprint]* doi: 10.1101/2020.02.26.966184
- Mukesh Sankar, S., Tara Satyavathi, C., Madan, P., Bharadwaj, C., Singh, S. P., and Barthakur, S. (2014). Genetic diversity analysis for high temperature stress tolerance in Pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Ind. J. Plant Physiol.* 19, 324–329. doi: 10.1007/s40502-014-0099-2
- Muthusamy, S. K., Dalal, M., Chinusamy, V., and Bansal, K. C. (2017). Genome-wide identification and analysis of biotic and abiotic stress regulation of small heat shock protein (HSP20) family genes in bread wheat. *J. Plant Physiol.* 211, 100–113. doi: 10.1016/j.jplph.2017.01.004
- Nagaraju, M., Reddy, P. S., Kumar, S. A., Kumar, A., Rajashekar, G., Rao, D. M., et al. (2020). Genome-wide identification and transcriptional profiling of small heat shock protein gene family under diverse abiotic stress conditions in *Sorghum bicolor* (L.). *Int. J. Biol. Macromol.* 142, 822–834. doi: 10.1016/j.ijbiomac.2019.10.023
- Nievol, C. C., Carvalho, C. P., Carvalho, V., and Rodrigues, E. (2017). Rapid responses of plants to temperature changes. *Temperature* 4, 371–405. doi: 10.1080/23328940.2017.1377812
- Ouyang, Y., Chen, J., Xie, W., Wang, L., and Zhang, Q. (2009). Comprehensive sequence and expression profile analysis of *Hsp20* gene family in rice. *Plant Mol. Biol.* 70, 341–357. doi: 10.1007/s11103-009-9477-y
- Pareek, A., Mishra, D., Rathi, D., Verma, J. K., Chakraborty, S., and Chakraborty, N. (2021). The small heat shock proteins, chaperonin 10, in plants: an evolutionary view and emerging functional diversity. *Environ. Exp. Bot.* 182:104323. doi: 10.1016/j.envexpbot.2020.104323
- Poidevin, L., Forment, J., Unal, D., and Ferrando, A. (in press). Transcriptome and translational changes in germinated pollen under heat stress uncover roles of transporter genes involved in pollen tube growth. *Plant Cell Environ.* 1–18. doi: 10.1111/pce.13972
- Reddy, P. S., Mishra, R., Chakradhar, T., Malik, S., Kaul, T., Sopory, et al. (2010). Molecular cloning and characterization of gene encoding cytoplasmic HSC 70 from *Pennisetum glaucum* may play a protective role against abiotic stresses. *Mol. Gen. Genomics* 283, 243–254. doi: 10.1007/s00438-010-0518-7
- Sairam, R. K., and Tyagi, A. (2004). Physiology and molecular biology of salinity stress tolerance in plants. *Curr. Sci.* 86, 407–421. doi: 10.1007/1-4020-4225-6
- Sajid, M., Rashid, B., Ali, Q., and Husnain, T. (2018). Mechanisms of heat sensing and responses in plants. It is not all about Ca²⁺ ions. *Biol. Plant* 62, 409–420. doi: 10.1007/s10535-018-0795-2
- Sali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815. doi: 10.1006/jmbi.1993.1626
- Sarkar, N. K., Kim, Y. K., and Grover, A. (2009). Rices Hsp genes: genomic organization and expression profiling under stress and development. *BMC Genomics* 10:393. doi: 10.1186/1471-2164-10-393
- Savic, J. A., Dragicevic, I., Pantelic, D., Oljaca, J., and Momcilovic, I. (2012). Expression of small heat shock proteins and heat tolerance in potato (*Solanum tuberosum* L.). *Arch. Biol. Sci. Belgrade* 64, 135–144. doi: 10.2298/ABS1201135S
- Scharf, K. D., Siddique, M., and Vierling, E. (2001). The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing a-crystallin domains (Acd proteins). *Cell Stress Chaperones* 6, 225–237. doi: 10.1379/1466-1268(2001)006<0225:tefoat>2.0.co;2
- Sharma, S., Sharma, R., Govindaraj, M., Mahala, R. S., Satyavathi, C. T., Srivastava, R. K., et al. (2021). Harnessing wild relatives of pearl millet for germplasm enhancement: challenges and opportunities. *Crop Sci.* 61, 177–200. doi: 10.1002/csc.2.20343
- Siddique, M., Gernhard, S., vonKoskull-Döring, P., Vierling, E., and Scharf, K. D. (2008). The plants HSP superfamily: five new members in *Arabidopsis thaliana* with unexpected properties. *Cell Stress Chaperones* 13, 183–197. doi: 10.1007/s12192-008-0032-6
- Singh, R. K., Jaishankar, J., Muthamilarasan, M., Shweta, S., Dangi, A., and Prasad, M. (2016). Genome-wide analysis of heat shock proteins in C₄ model, foxtail millet identifies potential candidates for crop improvement under abiotic stress. *Sci. Rep.* 6, 1–14. doi: 10.1038/srep32641
- Studer, S., Obrist, M., Lentze, N., and Narberhaus, F. (2002). A critical motif for oligomerization and chaperone activity of bacterial heat shock proteins. *Eur. J. Biochem.* 269, 3578–3586. doi: 10.1046/j.1432-1033.2002.03049.x
- Sun, M., Huang, D., Zhang, A., Khan, I., Yan, H., Wang, X., et al. and Huang, L. (2020). Transcriptome analysis of heat stress and drought stress in pearl millet based on Pacbio full-length transcriptome sequencing. *BMC Plant Biol.* 20, 1–15. doi: 10.1186/s12870-020-02530-0
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673
- Ullah, A., Salehnia, N., Kolsoumi, S., Ahmad, A., and Khaliq, T. (2018). Prediction of effective climate change indicators using statistical downscaling approach and impact assessment on pearl millet (*Pennisetum glaucum* L.) yield through genetic algorithm in Punjab, Pakistan. *Ecol. Indic.* 90, 569–576. doi: 10.1016/j.ecolind.2018.03.053
- Usman, M. G., Rafi, M. Y., Martini, M. Y., Yusuff, O. A., Ismail, M. R., and Miah, G. (2018). Introgression of heat shock protein (Hsp70 and sHsp) genes into the Malaysian elite chilli variety Kulai (*Capsicum annum* L.) through the application of marker-assisted backcrossing (MAB). *Cell Stress Chaperones* 23, 223–234. doi: 10.1007/s12192-017-0836-3

- van Montfort, R. L., Basha, E., Friedrich, K. L., Slingsby, C., and Vierling, E. (2001). Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat. Struct. Biol.* 8, 1025–1030. doi: 10.1038/nsb722
- Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35:10. doi: 10.1038/nbt.3943
- Wahid, A., Gelani, S., Ashraf and Foolad, M. R. (2007). Heat tolerance in plants: an overview. *Environ. Exp. Bot.* 61, 199–223. doi: 10.1016/j.envexpbot.2007.05.011
- Wang, J., Gao, X., Dong, J., Tian, X., Wang, J., and Palta, J. A., et al. (2020). Over-expression of the heat-responsive wheat gene *TaHSP23.9* in transgenic *Arabidopsis* conferred tolerance to heat and salt stress. *Front. Plant Sci.* 11:243. doi: 10.3389/fpls.2020.00243
- Waters, E. R., Lee, G. J., and Vierling, E. (1996). Evolution, structure and function of the small heat shock proteins in plants. *J. Exp. Bot.* 47, 325–338. doi: 10.1093/jxb/47.3.325
- Waters, E. R., and Vierling, E. (2020). Plant small heat shock proteins—evolutionary and functional diversity. *New Phytol.* 227, 24–37. doi: 10.1111/nph.16536
- Yang, J. Y., Zheng, W., Tian, Y., Wu, Y., and Zhou, D. W. (2011). Effects of various mixed salt-alkaline stresses on growth, photosynthesis, and photosynthetic pigment concentrations of *Medicago ruthenica* seedlings. *Photosynthetica* 49, 275–284. doi: 10.1007/s11099-011-0037-8
- Yao, F., Song, C., Wang, H., Song, S., Jiao, J., Wang, M., et al. (2020). Genome-wide characterization of the HSP20 gene family identifies potential members involved in temperature stress response in Apple. *Front. Genet.* 11:609184. doi: 10.3389/fgene.2020.609184
- Yu, J. H., Cheng, Y., Feng, K., Ruan, M. Y., Ye, Q. J., Wang, R. Q., et al. (2016). Genome-wide identification and expression profiling of tomato Hsp20 gene family in response to biotic and abiotic stresses. *Front. Plant Sci.* 7:1215. doi: 10.3389/fpls.2016.01215
- Zhang, J., Chen, H., Wang, H., Li, B., Yi, Y., Kong, F., et al. (2016). Constitutive expression of a tomato small heat shock protein gene *LeHSP21* improves tolerance to high-temperature stress by enhancing antioxidation capacity in tobacco. *Plant Mol. Biol. Rep.* 34, 399–409. doi: 10.1007/s11105-015-0925-3
- Zhao, P., Wang, D. D., Wang, R. Q., Kong, N. N., Zhang, C., Yang, C. H., et al. (2018). Genome-wide analysis of the potato Hsp20 gene family: identification, genomic organization, and expression profiles in response to heat stress. *BMC Genomics* 19:61. doi: 10.1186/s12864-018-4443-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Mukesh Sankar, Tara Satyavathi, Barthakur, Singh, Bharadwaj and Soumya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Development of Sorghum Genotypes for Improved Yield and Resistance to Grain Mold Using Population Breeding Approach

C. Aruna^{1*}, I. K. Das¹, P. Sanjana Reddy¹, R. B. Ghorade², A. R. Gulhane², V. V. Kalpande², S. T. Kajjiloni³, N. G. Hanamaratti³, S. N. Chattannavar³, Shivaji Mehtre⁴, Vikram Gholve⁴, K. R. Kamble⁴, C. Deepika¹, N. Kannababu¹, D. M. Bahadure¹, Mahalingam Govindaraj⁵ and V. A. Tonapi¹

OPEN ACCESS

Edited by:

Velu Govindan,
International Maize and Wheat
Improvement Center, Mexico

Reviewed by:

Divya Ambati,
ICAR- Indian Agricultural Research
Institute, India
Rahul Madhavrao Phuke,
Indian Agricultural Research Institute,
India

*Correspondence:

C. Aruna
aruna@millets.res.in

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 29 March 2021

Accepted: 24 June 2021

Published: 28 July 2021

Citation:

Aruna C, Das IK, Reddy PS,
Ghorade RB, Gulhane AR,
Kalpande VV, Kajjiloni ST,
Hanamaratti NG, Chattannavar SN,
Mehtre S, Gholve V, Kamble KR,
Deepika C, Kannababu N,
Bahadure DM, Govindaraj M and
Tonapi VA (2021) Development of
Sorghum Genotypes for Improved
Yield and Resistance to Grain Mold
Using Population Breeding Approach.
Front. Plant Sci. 12:687332.
doi: 10.3389/fpls.2021.687332

¹ ICAR-Indian Institute of Millets Research, Hyderabad, India, ² Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola, India, ³ UAS, Dharwad, India, ⁴ Vasantrao Naik Marathwada Krishi Vidyapeeth, Parbhani, India, ⁵ International Crops Research Institute for Semi-arid Tropics, Patancheru, India

The infection caused by grain mold in rainy season grown sorghum deteriorates the physical and chemical quality of the grain, which causes a reduction in grain size, blackening, and making them unfit for human consumption. Therefore, the breeding for grain mold resistance has become a necessity. Pedigree breeding has been widely used across the globe to tackle the problem of grain mold. In the present study, a population breeding approach was employed to develop genotypes resistant to grain mold. The complex genotype \times environment interactions (GEIs) make the task of identifying stable grain mold-resistant lines with good grain yield (GY) challenging. In this study, the performance of the 33 population breeding derivatives selected from the four-location evaluation of 150 genotypes in 2017 was in turn evaluated over four locations during the rainy season of 2018. The Genotype plus genotype-by-environment interaction (GGE) biplot analysis was used to analyze a significant GEI observed for GY, grain mold resistance, and all other associated traits. For GY, the location explained a higher proportion of variation (51.7%) while genotype (G) \times location (L) contributed to 21.9% and the genotype contributed to 11.2% of the total variation. For grain mold resistance, G \times L contributed to a higher proportion of variation (30.7%). A graphical biplot approach helped in identifying promising genotypes for GY and grain mold resistance. Among the test locations, Dharwad was an ideal location for both GY and grain mold resistance. The test locations were partitioned into three clusters for GY and two clusters for grain mold resistance through a “which-won-where” study. Best genotypes in each of these clusters were selected. The breeding for a specific cluster is suggested. Genotype-by-trait biplots indicated that GY is influenced by flowering time, 100-grain weight (HGW), and plant height (PH), whereas grain mold resistance is influenced by glume coverage and PH. Because GY and grain mold score were independent of each other, there is a scope to improve both yield and resistance together.

Keywords: grain mold, population breeding, GGE biplot, G \times E interactions, glume cover, grain hardness

INTRODUCTION

Sorghum is well-adapted to the harsh conditions of arid and semiarid agroecologies where other crops fail to thrive well. Sorghum grain has a wide range of uses such as food, feed, brewing, and grain-based ethanol production (Ashok Kumar et al., 2011; Gonzalez et al., 2011; Aruna et al., 2019). Though sorghum has an important role in different sectors, the availability of good quality grain in an adequate amount is greatly influenced by climate variability, which induces constraints in terms of biotic and abiotic stresses (Das et al., 2012; Tovignan et al., 2016). Changed climate scenario, cropping pattern, and the cultivation of high-yielding hybrids and varieties have brought a change in the disease scenario over time, and the diseases, which were considered minor at some point of time, have become predominant, thereby inducing losses in yield and quality. Grain mold of rainy season grown sorghum is one such disease, which has changed its significance over time (Das et al., 2020). Now, it has become a prominent disease of rainy season sorghum for its devastating effects on grain yield (GY) and quality. It is a common disease in the countries of Asia, Africa, North America, and South America (Frederiksen et al., 1982; Das and Padmaja, 2016). The severity of the disease has increased in Asia and Africa where white grain sorghum is grown for food use. In India, high-yielding white grain hybrids are widely grown during the rainy season, which are devastated due to grain mold. This is mainly due to the cultivation of short- and medium-duration sorghum cultivars that mature during the rainy days in humid, tropical, and subtropical climates. These cultivars become susceptible to the disease while late-maturing photoperiod-sensitive sorghums generally escape grain mold as they attain maturity during dry weather (Rao et al., 2013). The disease has adverse effects on GY, quality, market value, seed quality, and eventually on the end products developed from grain (Das, 2019). Mold attack has become one of the major constraints hindering sorghum cultivation, which resulted in a drastic decrease in the area of cultivation. GY reduction up to 30–100% was reported (Ashok Kumar et al., 2011) with the estimated annual loss of US\$ 130 million across the globe (ICRISAT, 1992) and US\$ 50–80 million in India (Das and Patil, 2013). Yield losses up to 100% are reported on grain mold-susceptible cultivars under conducive environmental conditions such as rains during grain maturity (Singh and Bandyopadhyay, 2000; Navi et al., 2005). The grain mold causes yield reductions by poor seed formation due to caryopsis abortion, reduced grain filling, and filling grains with less density. The grain quality and thereby the market value is affected due to surface discoloration, poor grain size, and density. In addition, grain mold poses health hazards to both human and animals due to the production of mycotoxins and secondary metabolites (Castor and Frederiksen, 1980; Little and Magill, 2009; Audilakshmi et al., 2011).

The prevailing weather conditions, particularly high relative humidity (RH) and temperature during the period between flowering and maturity, are the most important factors that influence grain mold development. Warm and humid weather is more conducive to the development of the disease. Humid weather during and after flowering is required for grain mold

development and longer the duration of such conditions, more is the incidence of mold development. The severity of fungal sporulation and grain mold increases on most sorghum lines with an increase in the incubation temperature from 25 to 28°C with RH levels of 95–98% (Tonapi et al., 2007). RH, rainfall, the number of rainy days, and minimum temperatures for 4–6 weeks after flowering are significantly correlated with mold incidence (Ratnadassa et al., 2003; Tarekegn et al., 2006).

Grain mold resistance is complex as it involves many fungi (Williams and Rao, 1981) and is governed by many traits (Ambekar et al., 2011). Several management strategies, such as fungicide application, chemical and physical treatments, and the adaptation of tolerant varieties, were tried to control grain mold infection (Navi et al., 2002; Indira et al., 2006). Among all methods, the use of resistant varieties is cost-effective and ecofriendly to control grain mold (Mofokeng et al., 2017). It has been reported by many authors that the resistance to grain mold is conferred by several physical and chemical properties of plant and grains, which include loose panicle architecture, red pericarp, hard grain with increased glume coverage, glume color, pigmented testa, phenolic compounds, high levels of condensed tannin, phenolic acids, flavan-4-ols, and antifungal proteins (Glueck and Rooney, 1980; Esole et al., 1993; Rodriguez-Herrera et al., 1999; Waniska et al., 2001; Audilakshmi et al., 2005; Ulaganathan, 2011; Little et al., 2012; Thirumala Rao et al., 2012). The effects of different plant characters, such as phenological traits (height and duration), panicle structure (compactness), and floral traits [glume cover (GC) and length], on grain mold resistance have been extensively explored. Grain mold resistance is under the control of polygenes involving both major and minor genes having additive and epistatic effects and significant genotype \times environment interactions (GEI; Klein et al., 2001; Audilakshmi et al., 2005). The involvement of a minimum of 4–10 genes was reported by Rodriguez-Herrera et al. (2000). Due to the complex nature of grain mold, its management has also been difficult.

A grain mold-resistant cultivar provides the best means for minimizing the yield losses due to the disease epidemic (Forbes et al., 1992; Prom and Erpelding, 2009; Prom et al., 2014). This approach is highly recommended in exploring host-plant resistance mechanism. However, most of the efforts in developing mold-resistant cultivars met with challenges to tackle with the complex nature of the fungal species involved, many mechanisms governing resistance, multiple genes involved in genetic resistance, and a significant impact of environment and thus could derive limited success in the development of mold-resistant genotypes (Prom et al., 2003; Audilakshmi et al., 2011; Mpofu and McLaren, 2014). Nevertheless, many efforts are being made to improve grain mold resistance in sorghum. Both conventional breeding approaches and molecular approaches have been followed to tackle the problem of grain mold. Mainly, pedigree breeding was followed by involving resistant sources in the crossing program and making selections in segregating populations. However, major breeding efforts in the last three to four decades to incorporate grain mold resistance in the high-yielding genetic background have not paid many dividends, the real challenge being the incorporation of different resistant

mechanisms into a single cultivar to make it mold resistant. Low-to-moderate effect quantitative trait loci (QTLs) conferring the resistance to grain mold were identified in QTL mapping studies (Klein et al., 2001). However, these QTLs did not have a major effect on grain mold resistance directly and were associated with the traits that indirectly modulate resistance. Recently, genome-wide association mapping (GWAS) studies identified a major grain mold-resistant locus containing tightly linked and sequence-related MYB transcription factor genes (Nida et al., 2019; Prom et al., 2020).

A population breeding approach was initiated in Indian Institute of Millets Research (IIMR) as one of the approaches to tackle the problem of grain mold. Population improvement methods provide an opportunity for obtaining desirable gene recombinants by breaking undesirable linkages between desirable and undesirable traits especially for the incorporation of biotic stress- and abiotic stress-resistant genes from agronomically non-elite germplasm lines into the elite background (Reddy et al., 2006). Population improvement involves two important steps. In the first step, broad genetic-based gene pools are created. Later on, they are further improved through recurrent selection methods, which involve a cyclic scheme that includes the selection of desired plants, and a selected recombination, which allows random mating among the selected plants, and the generation of different types of families/progenies from the selected plants. The superior ones are intercrossed for an increased recombination to constitute new populations. The repeated selection and crossing of desirable plants through the recurrent selection methods result in the release of concealed genetic variation due to increased opportunities for recombination, because of which the frequencies of favorable genotypes steadily increase in a population. Population improvement programs have been made feasible in sorghum through the discovery of genetic male sterility. The reciprocal recurrent selection methods are further used to exploit additive (A) and $A \times A$ and other epistatic gene actions (Comstock and Robinson, 1952; Eberhart, 1972). Many breeders have adopted the population improvement methods in sorghum (Doggett, 1972; Maunder, 1972). Of several sources of genetic male sterility, the *ms3* and *ms7* alleles have been extensively used in population improvement programs worldwide due to their stability across environments (Bhola, 1982; Reddy and Stenhouse, 1994; Murty and Rao, 1997). Of late, Bernardino et al. (2021) reported that a multiparental random mating population in sorghum could be used to detect QTLs related to tropical soil adaptation, fine mapping of underlying genes, and genomic selection approaches. Keeping the advantages of the population breeding in bringing together the desirable alleles, we initiated a population breeding approach in 2000 and the two populations were developed, one for female parent (B line) development and another for male parent (R) development. After completing three cycles of random mating and three cycles of half-sib family selections, the lines were stabilized. In the present study, 150 derivatives from a random mating population were screened for their grain mold resistance across the four locations. From these, 33 derivatives were selected and were evaluated for their yield traits and grain mold resistance over the four locations.

The GEI poses a challenge to the plant breeders during testing and the selection of desirable genotypes, resulting in slowing down of the genetic progress (Romagosa and Fox, 1993). A significant GEI may be the resultant of (i) a noncrossover GEI wherein the genotype rank does not significantly change across environments and a significant interaction is mainly due to the magnitude of the response of genotypes with a location, or (ii) a crossover type of GEI, where the rank of the tested genotype changes from one environment to another, which means that a genotype that is best performing in one environment may be a poor performer in another environment. Selection becomes easy for a plant breeder if the GEI is of non-crossover type while the crossover interactions pose difficulties in plant breeding (Mohammadi and Amri, 2013). In such situations, multi-environment trials (METs) help to identify the superior genotypes with broad as well as specific adaptation. The performance of genotypes across environments can be used to derive the relation among the testing locations and also group the locations into different mega-environments (MEs; Yan et al., 2000). Each ME consists of a group of testing locations that show a similar or non-crossover genotypic response as well as depict consistent performance of genotypes across years (Gauch and Zobel, 1997; Yan and Rajcan, 2002; Yan and Tinker, 2005). For interpreting the MET data, the GGE biplot developed by Yan et al. (2000) is preferred by researchers across the world as it allows a visual examination of the GEI pattern. The advantage of this method is that the genotype variance is integrated with the GEI effects removing the only variance due to E. The “which-won-where” view of the GGE biplot identifies the best performing genotype in a particular ME (Yan et al., 2007). In addition, for the most ideal genotype, best suitable environments can be visualized apart from a comparison of any two cultivars across environments, the determination of the yield, and the stability of genotypes in “mean vs. stability view.” An ideal environment for initial testing can be identified by looking into discriminability and representativeness of the environments apart from grouping the locations into MEs (Malla et al., 2010).

The association among the traits across locations can be viewed from the genotype-trait (GT) biplot and an application of the GGE biplot technique where traits are considered as testers (Yan and Rajcan, 2002; Yan and Tinker, 2005; Dehghani et al., 2008). The GT biplots summarize the genotype-by-trait matrix graphically (Yan and Rajcan, 2002; Yan and Kang, 2003). And the graph can also identify the specific trait combinations to be employed for indirect selection (Yan and Tinker, 2005). In addition, the potential known genotypes that are good performers for specific traits give insights into a parental line with trait combinations for enhancing yield. The GGE biplot technique was used to understand the stability of genotypes and also to interpret the complexity of GEI in grain sorghum (Rakshit et al., 2012), forage sorghum (Aruna et al., 2016), and sweet sorghum (Rao et al., 2011).

The present study aims to evaluate the performance and stability of 33 derivatives from the population improvement program for grain mold resistance and GY across the four locations. The main objectives were (i) to interpret the complexity of GEI in GM using the GGE biplot analysis; (ii) to

identify stable lines with the grain mold resistance that could be utilized in the resistance breeding programs; and (iii) to understand trait associations among grain mold traits using the GT biplot analysis.

MATERIALS AND METHODS

Plant Material

The plant material consists of the derivatives of the population breeding approach. In the early 2000s, population breeding has been initiated at ICAR-IIMR to facilitate the accumulation of genes from different resistance sources to tackle the grain mold problem using a half-sib mating system. Two separate populations were maintained, one for B line and another for R line improvement. Care was taken to involve a different set of lines in each of the populations to maintain genetic diversity. In the B population, elite B lines (296B, 27B, 422B, and 463B), some elite sources of grain mold resistance (SPGM 950267, SPGM 950283, and SPGM 950288), and grain mold-resistant sources from germplasm (B 58586 and IS 25017) were used. In the R population, elite R lines (C43, C85, NR 9, and NR 486), improved the grain mold-resistant source (SR 839) and the resistant source from germplasm (IS 14332), were involved. Three cycles of random mating were allowed in each of the populations in isolation. Later, three cycles of selections through half-sib mating were carried out. Each cycle of half-sib mating consists of one generation of selection for the target trait and one generation of random mating among the selected lines, thus increasing the frequency of desirable alleles. After three cycles of half-sib mating, the population mean of grain mold had come down to 5.6 score (on the scale of 1–9) from an initial score of 8.0 in case of the B line population, whereas it had come down to 5.4 from an initial score of 7.0 in case of the R line population. After three cycles of half-sib mating, the lines were made genetically stable by selfing over 3–4 generations. A total of 150 such stable lines were evaluated in the four locations during 2017. Out of these, 33 derivatives that performed well across the locations were selected and again tested in the locations during 2018.

Experimental Location and Design

The trials were conducted in the four testing locations during 2017 and 2018. The chosen locations were from sorghum growing states such as Maharashtra, Karnataka, and Telangana. These locations were distributed across the three states of India, with two locations (Akola and Parbhani) in Maharashtra, one each in Karnataka (Dharwad) and Telangana (Hyderabad). All these four locations were earlier identified as hot-spot locations for sorghum grain mold under All India Co-ordinated Research Project on Sorghum (AICRP-Sorghum). The climatic conditions of the selected locations for this study are presented in **Figure 1**.

In 2017, 150 random mating-derived lines were sown in an augmented design in the four locations. These lines were shown in three blocks each with 50 lines. In each block, five checks were included. The checks include two high-yielding but grain mold-susceptible B lines (296B and 27B), two moderately grain mold-resistant R lines (C43 and CB33), and one grain mold-resistant check (B 58586). In 2018, the experimental design was

a randomized complete block (RCBD) with three replications. Two checks, one grain mold-resistant check, B 58586 and an elite grain mold-susceptible check, 296B were used for the experiment. Each genotype had a plot size of two rows of 4 m length each, with 0.6 m between rows and 0.15 m between hills in each row. In each location, the experiment was planted with the onset of rains during June. All the recommended crop production and protection practices were followed to raise a crop with good plant stand. Standard crop management practices were followed across all locations.

Field Monitoring and Data Collection

In 2017, observations on grain mold score, days to flower (DF), plant height (PH), GY per plant, and 100-grain weight (HGW) were recorded, whereas, in 2018, all the following observations were recorded in each location.

Phenological and Yield Traits

Days to Flower

Number of days to anthesis for ~50% of the plants to reach mid bloom was recorded for each plot.

Plant Height

PH was recorded at physiological maturity on 10 plants per plot, from the soil surface to the top of the panicle and was expressed in cm.

Panicle Length

The length of the panicle from the base to the tip was measured on 10 plants per plot at the time of harvest and was expressed in cm.

Grain Yield

The grain of 10 plants was harvested individually and weighted. The average weight of the 10 plants was taken as the yield per plant in grams. The first and last plants in each row were not harvested to eliminate the confounding results caused by a border effect.

Hundred-Grain Weight

The weight of 100 grains from each panicle was recorded and expressed in grams. The average of 10 plants is used for analysis.

Grain Hardness

Digital hardness tester (Pharmag Instruments Ltd., Kutkatpally, Hyderabad, India) was used to record grain hardness (GHR) for the grain samples from Hyderabad and Akola. It employed a force gauge to measure tension/compression force applied on an individual seed. The measurement on 10 randomly selected seeds per plant and on 5 plants per replication was taken. The individual seed hardness from each sample was tested, and the mean value was expressed in kilogram force (kgf).

Grain Mold Score

Threshing of grains were carried out plot-wise, and the grain mold reactions of each plot were measured as grain mold score (GMS) as per Audilakshmi et al. (2011). About 20 g of the grain sample was placed in a petri plate, and grain mold infection

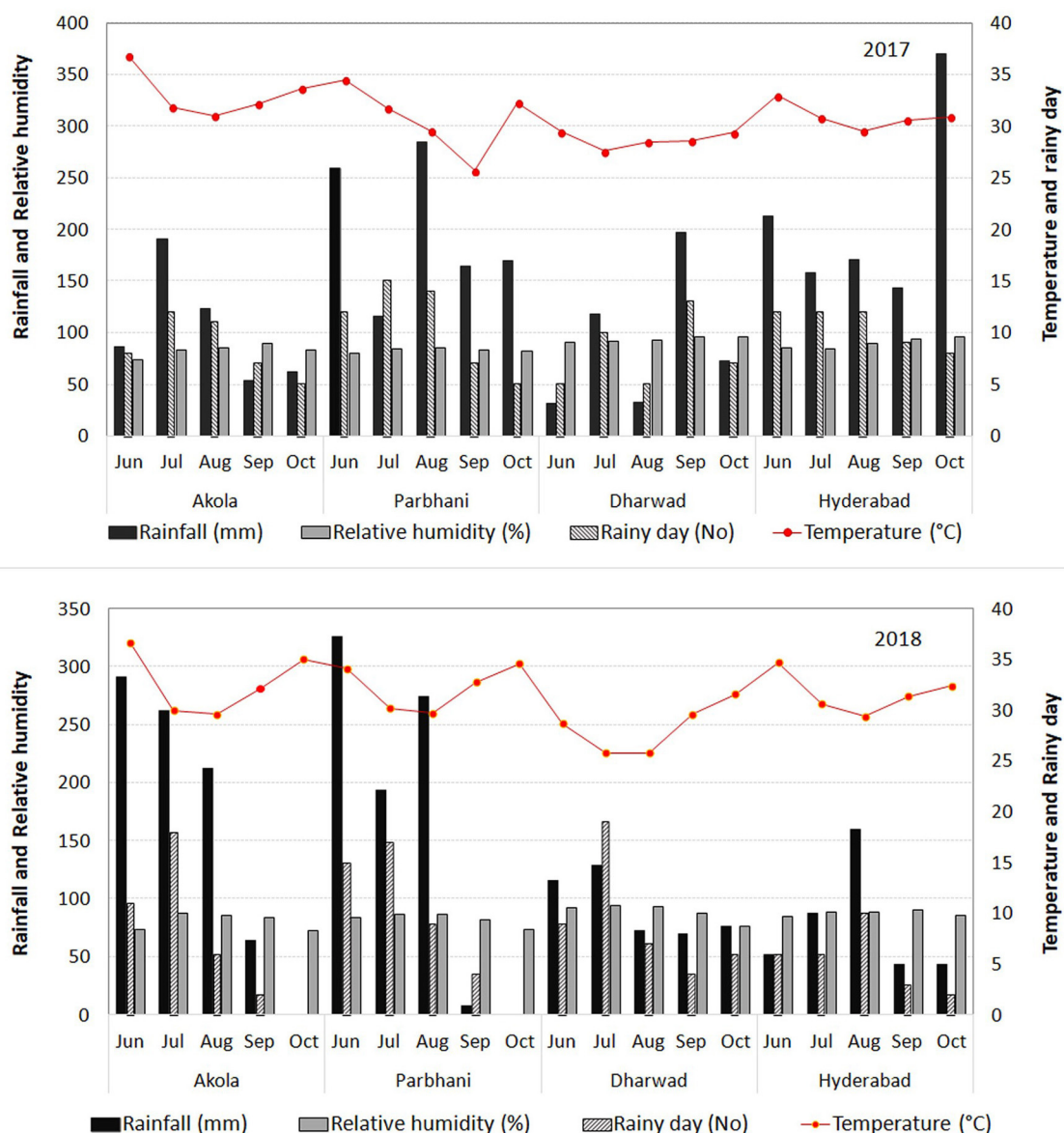


FIGURE 1 | Meteorological data [rainfall, the number of rainy days, relative humidity (RH), and temperature] in the four locations of the study during 2017 and 2018.

was scored visually using a 1–9 scale, where 1 = no mold and 9 = more than 75% of grains in the sample are with mold. The GMS data were subjected to square root transformation for further analysis.

Glume Cover

GC was assessed visually based on the portion of the grain covered with glume as 25, 50, 75, and 100%. This visual assessment was taken on 10 plants in each plot and the average was drawn. The data were transformed using an arcsine transformation.

Glume Color

GCL was recorded visually as white, red, and brown on 10 plants in each plot.

Panicle Compactness

Panicle compactness (PC) is recorded visually as loose, semi-compact, and compact based on the density of the ear head at the time of maturity.

Data Analysis

The 2017 data of all the four locations were subjected to the individual and pooled augmented analysis through the augmented complete block design (ACBD) in R program

(Rodríguez et al., 2018). Out of 150 genotypes tested, 33 superior derivatives were selected for evaluation in 2018 based on their grain mold score across the locations in 2017. For the data of the four locations in 2018, ANOVA was carried out to test location (L), genotype (G), and their interaction effects ($G \times L$). Trait variability and correlations were calculated using replicate means. The replication-wise data were analyzed for ANOVA using Genstat 12th edition (V. S. N., International, 2009). Heritability (h^2b) is calculated for the target traits as the ratio of genetic variance to the phenotypic variance (which is the total of the genetic variance and environmental variance). The levels of the broad-sense heritability (h^2bs) are categorized as low (<0.3), moderate ($0.3-0.6$), and high (>0.6) according to Robinson et al. (1949).

The GGE biplot analysis was employed to interpret the genotype by environment interactions for GY and grain mold tolerance. A GGE biplot methodology, which explains two concepts, the biplot concept (Gabriel, 1971) and the GGE concept (Yan et al., 2000), was used to analyze the data. The statistical theory of GGE methodology as given by Yan and Kang (2003) was considered. The data were analyzed as described by Rakshit et al. (2012) using the software GGE biplot ver. 8.2 (Yan, 2001). The MLT data were analyzed in the GGE biplot with the options (scaling = “scaling 4”) and tester-centered (centering = 2) GGE biplot, as suggested by Yan and Holland (2010). For obtaining genotypes with high yield and stability using the “mean vs. stability” option, genotype-focused singular value partitioning (SVP = 1) was used, and for evaluating locations, environment-focused SVP = 2 was employed (Yan, 2001) using the “relation among testers” option. The “which-won-where” option was used to pool the locations into MEs and to identify the promising genotype in a given ME.

Genotype-By-Trait Biplot

The GGE biplot software was used to generate genotype \times trait biplots by the “genotype-by-trait biplots” option using the pooled data and “scaling = 1.” Traits were considered as testers. Phenotypic correlations among traits were determined using trait-focused SVP (Yan and Tinker, 2005).

RESULTS

ANOVA and Mean Performance

The mean values and ranges of trait expression of the studied 150 genotypes in 2017 are presented as location wise in **Table 1**. The grain mold incidence was high in Hyderabad and Dharwad as revealed by the location mean and range of grain mold score. Good amount of variation was observed for all other traits. A total of 33 derivatives were selected based on their better level of grain mold resistance across the four locations for further study during 2018. The results of ANOVA and the mean values for all the traits studied during 2018 are presented in **Tables 2, 3**. The ANOVA showed that the effects of G and L and $G \times L$ interaction effects were significant ($p < 0.01$) for all traits. The DF ranged from 66 to 80 days with G6, G7, G8, G5, G15, and G13 flowering in <70 days. It was longer in Akola, whereas it was shorter in Hyderabad. There was a significant effect of G, L, and the interaction G

TABLE 1 | Mean performance of all genotypes at different locations during the rainy season of 2017.

Centre	Trait	Range	Mean	Lsd (5%)
Akola	DF	57–92	70.1	7.82
	GMS	1.8–7.74	4.0	1.71
	PH	93.9–343.5	163.8	15.59
	HGW	1.04–3.31	2.35	0.37
	GY	8.87–45.9	22.26	12.13
Parbhani	DF	53–73	59	5.05
	GMS	1.93–8.27	4.88	1.66
	PH	96–403	191.9	0.004
	HGW	1.01–5.20	2.01	0.68
	GY	6.73–88.5	52.9	22.17
Dharwad	DF	57–74	67.0	2.86
	GMS	2.74–8.24	5.87	2.12
	PH	81.7–323.3	171.8	49.6
	HGW	1.50–3.59	2.24	0.50
	GY	9.94–62.9	35.34	14.36
Hyderabad	DF	54–81	70.2	6.66
	GMS	2.30–9.16	6.8	1.67
	PH	98.5–326.1	167.7	29.6
	HGW	1.27–3.73	2.37	0.52
	GY	8.91–87.7	39.59	14.68
Across	DF	63–76	68.8	4.29
	GMS	2.99–7.55	5.56	1.32
	PH	119.4–270.3	173.8	38.83
	HGW	1.79–3.13	2.24	0.38
	GY	25.62–49.74	37.55	11.84

DF, Days to flower; PH, Plant height (cm); HGW, 100-grain weight (g); GY, Grain yield (g/plant); GMS, Grain mold score.

\times L on flowering. The genotypes varied significantly for PH, which ranged from 123 to 223 cm. PH observed at Dharwad and Hyderabad was higher compared to that observed at Akola and Parbhani. For GY also, the differences among the genotypes were significant and there are a few genotypes, such as G28, G21, G10, G23, G27, and G1, that showed significantly higher performance for GY per plant over the elite check, 296B. In general, GY was less at Akola and Parbhani. The lines, G15, G28, and G7 exhibited significantly higher HGW over the elite check, 296B. The differences in grain weight among the locations was not much. Most genotypes recorded higher grain mold score in 2017 compared to 2018. For this trait, the genotypes G4, G33, G31, and G8 were statistically on par with the resistant check, B 58586 in both the years. However, all other test lines were significantly better than the elite check, 296B. Their GMS ranged from 2.78 to 5.54 in 2018, which indicates that the lines were all tolerant. The elite parent 296B had a GMS of 8.5 in 2017 and 7.75 in 2018 indicating that it is susceptible. The highest GMS among the locations was recorded at Hyderabad (6.8) followed by Dharwad (5.87) in both years. Though there were significant differences among locations, all the genotypes were categorized as tolerant based on the observed GMS values at these locations. GHR, which was measured at two locations, ranged from 6.33 to

TABLE 2 | (A) ANOVA for different studied traits across four locations during the rainy season of 2018. **(B)** ANOVA for grain yield and grain mold score in individual locations during rainy season of 2018.

(A)										
	df	DF	PH	GC	PL	GY	HGW	GMS	df	GH
Genotype	34	126.2**	7538.7**	1919.75**	74.21**	912.9**	0.63**	6.56**	34	2.28**
Location	3	2113.67**	70544.1**	2488.41**	206.9**	47670**	4.62**	31.98**	1	660.4**
Genotype × Location	102	34.26**	1588.1**	575.34**	25.43*	594.7**	0.36**	2.89**	34	2.20**
Residual	278	14.87	101.1	85.93	18.17	150	0.09	1.241	138	0.04
(B)										
Source	Grain yield					Grain mold score				
	df	Akola	Dharwad	Hyderabad	Parbhani	Akola	Dharwad	Hyderabad	Parbhani	
Replication	2	79.4	558.9	14.65	41.3	35.3	8.15	0.43	0.44	
Genotype	34	123.2**	1907.2**	500.2**	166.3*	2.01**	7.17**	3.78**	1.54*	
Residual	68	25.1	384.9	46.0	145.4	0.42	3.00	0.64	1.24	

* $p < 0.05$; ** $p < 0.01$; DF, Days to flower; PH, Plant height; GC, Glume cover; PL, Panicle length; GY, Grain yield; HGW, 100-grain weight; GMS, Grain mold score; GHR, Grain hardness.

12.33 kgf. The grain samples of Akola were harder compared to those of Hyderabad.

Broad-Sense Heritability

The h^2 bs for the yield and grain mold-resistant traits for an individual location and across the locations is presented in **Table 4**. Although the observed h^2 b was generally high for most of the traits, it showed a variation for each trait in respective locations. h^2 b was high for days to flowering, PH, GC, and HGW in all locations and across locations. GY had high h^2 b at Akola (0.8), Dharwad (0.8), and Hyderabad (0.91), whereas it had low h^2 b at Parbhani (0.13). Across locations, it was found to have moderate h^2 b (0.35). Grain mold score had moderate h^2 b at Akola (0.53) and Dharwad (0.6), whereas it showed high h^2 b at Hyderabad (0.83). Parbhani recorded low h^2 b for grain mold score (0.19). GHR recorded moderate h^2 b in both the locations where it was recorded.

Contribution of the Factors to the Variation of GY and Grain Mold Score

The relative contribution of variation due to genotype, environment, and $G \times E$ to the total variance measured for GY and GMS are presented in **Table 5**. The location with 51.7% of the total variance was the major contributor toward a variation in GY followed by the $G \times L$ interaction (21.9%). The genotype accounted for 11.2% of total variability. Similarly, a variation in GMS was largely explained by $G \times L$ (30.7%) followed by the G (23.3%) and L (10%).

GGE Biplot Analysis

For ease of interpretation of a graphical representation, grain mold score was rearranged with scaling 9 = resistant and 1 = susceptible for the GGE biplot analysis. Because the heritability was different in different environments, the GGE biplot analysis was done by applying h^2 b-adjusted GGE as suggested by Yan and Holland (2010).

Mean Performance and Stability Analysis of Genotypes Across Locations

The performance of genotypes for mean and stability is depicted by the average environment coordination (AEC) method and is presented in **Figure 2** (Yan, 2001, 2002). The environment-centered (centering = 2) genotype-metric (SVP = 1) biplots for each of the traits are presented in **Figures 2A–G**. The first two PCs explained about 94.7% of a variation for GY. In **Figure 2** the AEC abscissa, a line with a single arrow head, passes through the biplot origin and points toward the higher mean values. The perpendicular lines to the AEC abscissa are the AEC ordinates and their length indicates stability. The genotype, which is present away from the origin toward the arrow head, has the highest and in the opposite direction has the lowest mean GY. Accordingly, genotype G28 which is placed on the positive side at the far end from the origin, is the highest yielding, followed by G21. On the other hand, G33 had the lowest GY, followed by G7. The biplot results were compared with those of tabulated values for mean GYs of the genotypes (**Table 3**), thereby showing that the biplot can be used to visualize ranking of genotypes. The projection of the genotypes on the AEC ordinate approximated the stability with a greater projection indicating toward instability. The genotype G28 was less stable for GY with a higher projection from the AEC abscissa. On the contrary, G21 was relatively more stable as it was closer to the AEC abscissa. It also had the highest GY after G28.

For grain mold score, the first two PCs explained 83.8% of the variation. Here, the resistant genotype (G34) had the best level of resistance, being projected farthest from the origin. The genotypes, G4 and G33 had good levels of grain mold resistance and were more stable as their projection on the AEC ordinate is less. Though G8 and G9 also had low grain mold scores, they are not very stable.

Grain yield and grain mold resistance can be simultaneously improved indirectly through other target traits. The mean and stability biplots for all these traits are presented in **Figures 2C–G**. The analysis depicts desirable genotypes for each of these traits.

TABLE 3 | *Per se* performance of genotypes for all traits across locations during the rainy seasons of 2017 and 2018.

Genotype	DF	PH (cm)	GC (%)	PL (cm)	GY (g/pl)	HGW (g)	GMS		GH (kgf)
							2017	2018	
G1	78	189	67	31.5	53.06	2.50	4.18	4.43	12.33
G2	80	123	49	23.1	34.58	2.06	4.74	4.58	7.83
G3	78	144	68	31.3	44.48	1.90	5.09	3.49	7.67
G4	72	198	61	27.6	35.29	2.30	3.95	2.78	7.83
G5	68	172	41	23.1	39.71	2.34	5.18	3.80	8.5
G6	66	163	61	25.9	34.92	2.42	5.41	3.93	10.17
G7	67	139	55	24.0	32.93	2.44	5.70	5.37	6.83
G8	67	166	50	29.3	52.35	2.59	3.67	3.40	9.83
G9	77	170	49	27.6	42.50	2.38	4.03	3.43	9.17
G10	72	207	49	25.5	58.08	2.13	5.14	4.35	6.33
G11	77	174	48	27.9	45.94	2.28	4.54	4.82	7.17
G12	73	223	44	29.5	46.96	2.24	4.05	3.47	9.67
G13	70	148	52	24.4	37.19	2.24	5.97	4.32	7.33
G14	72	192	68	31.2	43.67	2.17	5.18	4.66	8.33
G15	69	183	52	26.7	37.00	2.71	5.61	4.57	10.33
G16	74	141	54	26.9	47.31	2.35	5.21	4.19	9.17
G17	76	132	35	27.7	43.73	2.47	5.34	5.03	8.67
G18	74	155	41	27.0	40.49	2.31	5.61	4.35	7.83
G19	80	154	30	27.5	45.17	2.43	4.68	4.78	8.67
G20	71	179	31	23.2	37.08	2.34	4.61	4.24	9.17
G21	76	152	70	23.7	61.29	2.36	4.71	4.57	9.33
G22	79	182	48	28.2	46.94	2.15	5.24	4.33	8.0
G23	77	174	57	25.1	55.63	2.36	4.12	3.83	7.33
G24	77	169	28	22.2	47.77	1.71	5.21	4.34	8.33
G25	77	180	45	27.0	49.88	1.89	4.14	3.61	9.17
G26	73	221	50	26.8	47.77	2.41	3.91	3.74	7.33
G27	79	198	56	23.5	53.83	2.36	3.85	3.56	9.17
G28	80	191	51	23.1	71.58	2.69	4.74	4.17	7.17
G29	75	145	42	26.3	33.98	2.21	3.75	3.87	8.33
G30	79	159	45	25.7	49.94	2.38	5.41	5.54	9.17
G31	75	150	51	26.7	48.52	2.39	4.06	3.31	8.33
G32	74	158	57	22.8	48.55	2.53	4.43	4.11	10.17
G33	76	156	65	26.8	31.90	2.02	3.50	2.93	9.83
G34 (B 58586)	79	202	90	27.6	38.88	1.78	2.89	2.52	9.67
G35 (296B)	79	134	36	27.8	42.75	2.35	8.50	7.75	7.83
Location									
Akola	80.4	154	56	25.0	25.29	2.17	4.0	3.43	8.63
Dharward	75.4	192	51	25.5	66.82	2.43	5.87	4.20	-
Hyderabad	69.4	190	45	28.1	60.08	2.50	6.8	4.77	6.08
Parbhani	75.5	141	53	27.0	28.57	2.25	4.88	4.08	-
Grand mean	75	169	51	26.4	45.19	2.29	5.39	4.12	6.86
Lsd (5%)	3.10	8.08	7.45	3.40	9.84	0.24	1.32	0.90	0.23
CV	5.1	5.9	18.1	16.1	27.1	13.0	-	27.0	2.9

The first two PCs explained 76.2–96.1% of the total variation in the data for different traits, thereby these graphs can be used for the interpretation of MLT data. For days to flowering, PC1 and PC2 explained 79.7% of the total variation and G6 was the earliest genotype with 66 DF, followed by G7 and G5. For PH, the biplot explained 96.1% of the variation. The genotypes G12 and G26 are

the tallest and G2 is the shortest. The biplot of panicle length (PL) explained 90.9% of the total variation and G3 and G14 had longer panicles. For HGW, the biplot explained 79.8% of the variation and the genotypes, G28 and G15 had a bold seed with the grain weight of around 2.7 g/100 grain. The biplot of glume coverage explained 76.2% of the total variation. The longest glumes were

TABLE 4 | Broad-sense heritability (h^2 bs) for the traits location wise and across locations during the rainy season of 2018.

Trait/Location	Akola	Dharwad	Hyderabad	Parbhani	Pooled data
Days flower	0.98	0.98	0.88	NE**	0.73
Plant height	0.97	0.99	0.97	0.91	0.79
Glume cover	0.80	0.82	0.93	0.84	0.7
Panicle length	0.76	0.44	0.89	0.03	0.66
Grain yield per plant	0.8	0.8	0.91	0.13	0.35
100 grain weight	0.97	0.81	0.86	0.75	0.63
Grain mold score	0.53	0.6	0.83	0.19	0.56
Grain hardness	0.58	NA*	0.38	NA*	0.40

*NA indicates trait values are not measured in this year for this location.

**NE indicated that heritability (h^2 b) is not calculated because their variance is negative.

TABLE 5 | Sum of squares and proportion of variation explained by year (Y), location (L), genotype (G), and their interactions on grain yield (GY) and grain mold score (GMS).

	df	Grain yield		GMS	
		Sum of squares	%ssT	Sum of squares	%ssT
Genotype	34	31038.3	11.2	223.1	23.3
Location	3	143009.9	51.7	95.9	10.0
Genotype \times Location	102	60655.4	21.9	294.5	30.7
Residual	278	41688.4	15.1	345.0	36.0
Total		276392		958.6	

observed in the resistant genotype (G34). Among the test lines, G21, G3, and G14 had longer glumes (68–70%). The numerical values of these traits (**Table 3**) were correlated with the graphical results, except a few minor changes in genotype ranking.

Because GMS and GY are the economically important traits, these traits are considered for further analysis. An “ideal genotype” is defined as one, which has high mean yield and stability across locations. Ranking of the other genotypes in comparison to the “ideal genotype” for GY is presented in **Figure 3** (Yan and Tinker, 2006). In the present study, genotypes G28, G21, G27, and G10 were closer to the ideal genotype (**Figure 3A**) and also had high numerical GY among the tested genotypes (**Table 3**). For grain mold resistance, G34, G4, and G33 were closer to an ideal one indicating they had a better level of grain mold tolerance. G28 was the highest yielder and was selected to understand the specific adaptation by ranking the test environment based on the relative GY of this genotype in the given environment (**Figure 4A**). Genotype 28 had the highest yields at Dharwad, above average performance at Hyderabad, near average at Akola, and below average at Parbhani. For grain mold resistance, G34 had a high level of resistance at Dharwad, followed by Hyderabad (**Figure 4B**). Its performance in all other locations was good and above average.

Ideal Test Locations for GY and Grain Mold Resistance

The MLT data can also be used to derive the information about testing locations (Cooper et al., 1996; Yan et al., 2007).

Graphical representation of the test environment can be viewed from the GGE biplot. An ideal environment is one, which can differentiate among the genotypes that can be measured by its vector length in the biplot, and it should represent all other locations which is, measured by its angle with the “average environment” (Yan, 2001). In the GGE biplot, the relationships among test environments can be derived by environment-metric (SVP = 2). The cosine of the angle between the two environments is related to the similarity between them. An acute angle indicates a closer relationship between the environments in ranking the genotypes, whereas an obtuse angle is indicative of a negative correlation and a right angle of no relation (Yan and Tinker, 2006). Combined analysis of the 35 genotypes for GY (**Figure 5ai**) showed that the locations Hyderabad, Dharwad, and Akola were positively correlated, whereas Parbhani with an obtuse angle with other locations had a negative association. For GMS, Dharwad, Parbhani, and Hyderabad were positively correlated with an acute angle, whereas Akola had no linear correlation with Dharwad and Parbhani but positively correlated with Hyderabad (**Figure 5aii**). The results of correlations among environments using the GGE biplot for GY and grain mold score were further confirmed by carrying out rank correlations among the environments (**Table 6**).

The discriminating ability and representativeness of the environments are visualized in **Figure 5bi,ii** for GY and grain mold resistance, respectively. The concentric circles on the biplot are proportional to the SD of the environments. It indicates the discriminating ability of the environments (Yan

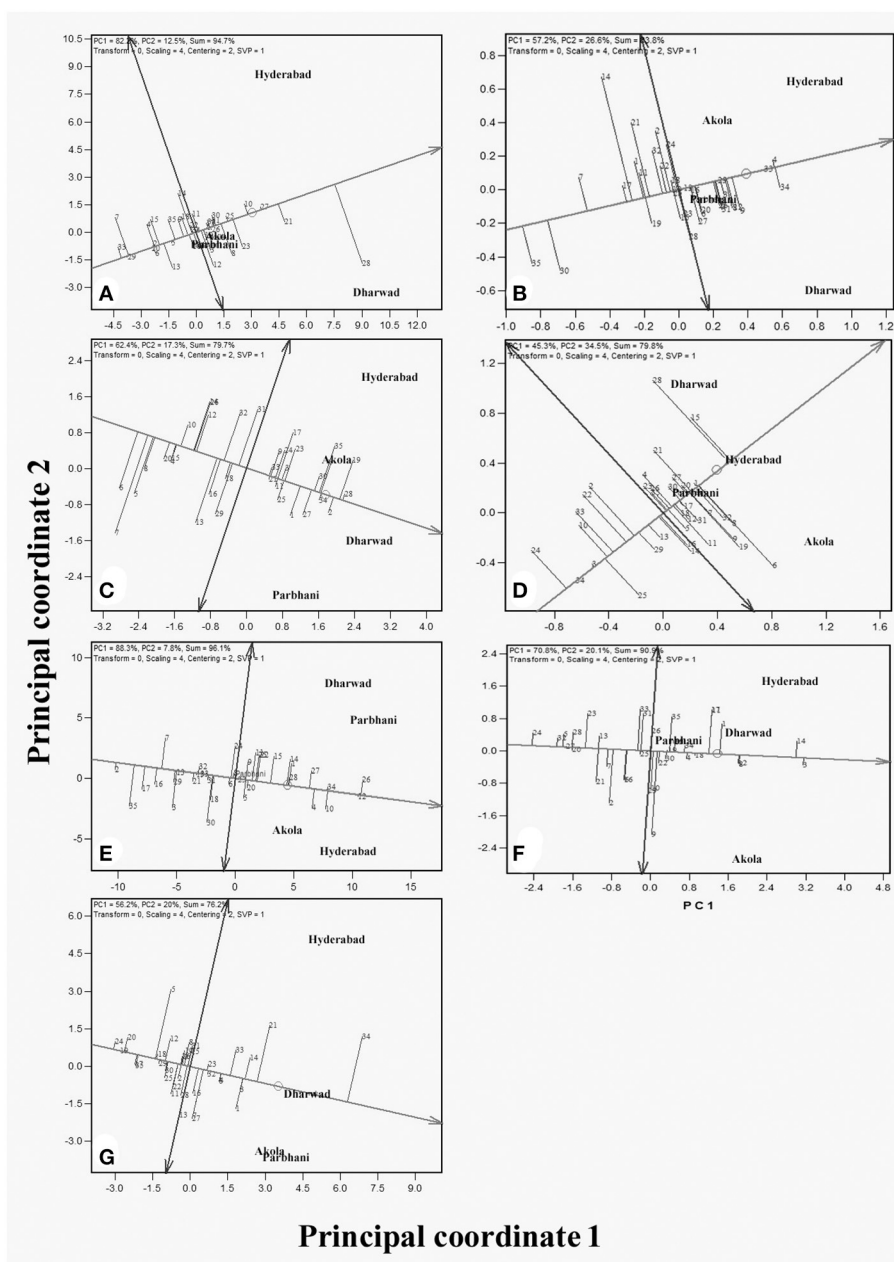


FIGURE 2 | GGE biplots from the combined analysis of data from 35 genotypes (1–35) evaluated across four locations for the traits: **(A)** grain yield (GY), **(B)** grain mold score, **(C)** flowering time, **(D)** 100-grain weight (HGW), **(E)** plant height (PH), **(F)** panicle length (PL), and **(G)** glume coverage.

and Tinker, 2006). In addition, the length of the environment vector is directly related to the discriminating ability. Dharwad with the longest vector length was found to be the most discriminating location, followed by Hyderabad for both GY and grain mold resistance. Parbhani and Akola revealed no discriminating power as depicted by a shorter vector length, and the conclusions on the genotypic performance may not be meaningful. Representativeness of a location can be obtained by projecting the locations with respect to the AEC abscissa (Figure 5b). The “average environment” is represented by a small

circle on the AEA in Figure 5b. Those environments, which make smaller angles with the AEA, are the most representative of all the test environments. Thus, Dharwad was closest to the average environment in case of GY, whereas both Hyderabad and Dharwad had almost the same angle in case of grain mold score. An “ideal” location is the one, which should be discriminating as well as represent the test locations (Yan, 2001). Thus, Dharwad emerged as an ideal location for the selection of genotypes with high yield and adaptability. For ranking the genotypes for GY in the near average environment, Dharwad (Figure 5ci) showed

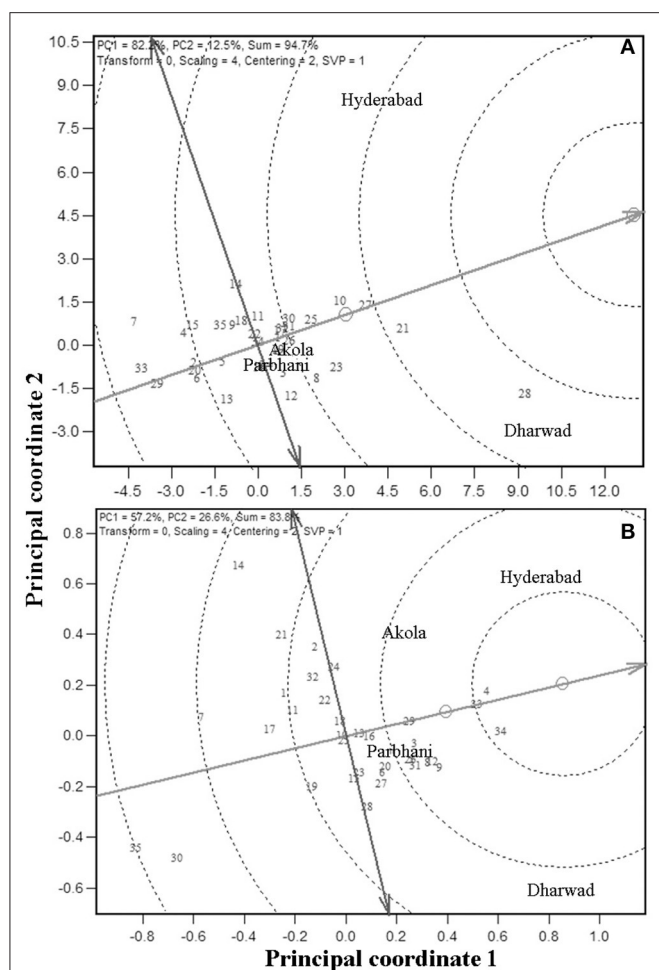


FIGURE 3 | Ranking of 35 genotypes relative to ideal genotype [the small circle on average environment coordinate, average environment coordination (AEC)] for GY (A) and GMS (B).

that at Dharwad, G28 yielded maximum followed by G21, G27, and G23, whereas G7 followed by G33 recorded the lowest yields. The genotypes, G16 and G22 showed near average GYs. For grain mold resistance at Hyderabad, G4, G33, and G34 showed high levels of resistance, whereas G35 and G30 recorded lower levels of resistance (Figure 5ci).

“Which-Won-Where” Analysis and Clustering of Environments

“Which-won-where” feature of the GGE biplot is exploited by researchers in many crops to derive the best performing genotypes in a subset of environments. In this view, the genotypes that are far away from the biplot origin are joined to form a polygon, and perpendicular lines are drawn from the origin to each side of the polygon. This splits the polygon into several sectors with a genotype at the vertex. The “Which-won-where” biplot for GY is presented in Figure 6A. The biplot indicated the existence of crossover GEIs and clusters of environments. The polygon had seven genotypes, viz., G28, G27, G14, G7, G33, G29,

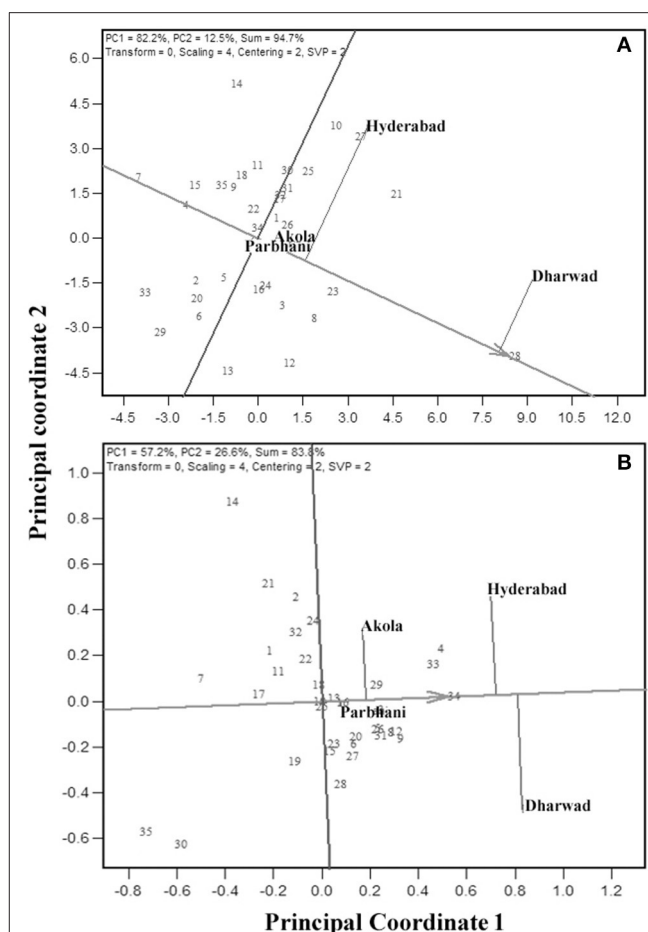


FIGURE 4 | Ranking of four environments based on the performance of the highest yielding genotype, 28 for GY (A) and most resistant genotype, 34 for grain mold score (B).

and G13 at the vertices. The perpendicular lines divided the biplot into seven sectors, of which three of them included all the four locations. Thus, the test locations were divided into three clusters, one with Dharwad and Akola with G28 as the best performing genotype for GY. The second cluster encompassed Hyderabad with G27 as the best genotype. The third cluster had Parbhani with G13 as the best genotype. For GMS, six genotypes, viz., G34, G28, G30, G35, G14, and G4 are placed at the vertices (Figure 6B) of the hexagon. Of the six sectors, two of them included four locations. Thus, the test locations were divided into two clusters. One with Dharwad and Parbhani with G34 as the best genotype. The second cluster had Hyderabad and Akola with G4 as the best genotype.

Trait Associations Through Genotype-By-Trait Biplot

A visualization of the relationship among the studied seven traits is facilitated by the $G \times T$ plot, an extension of GGE biplot where the traits are considered as testers (Figure 7). The association between any two traits is represented by the cosine of the

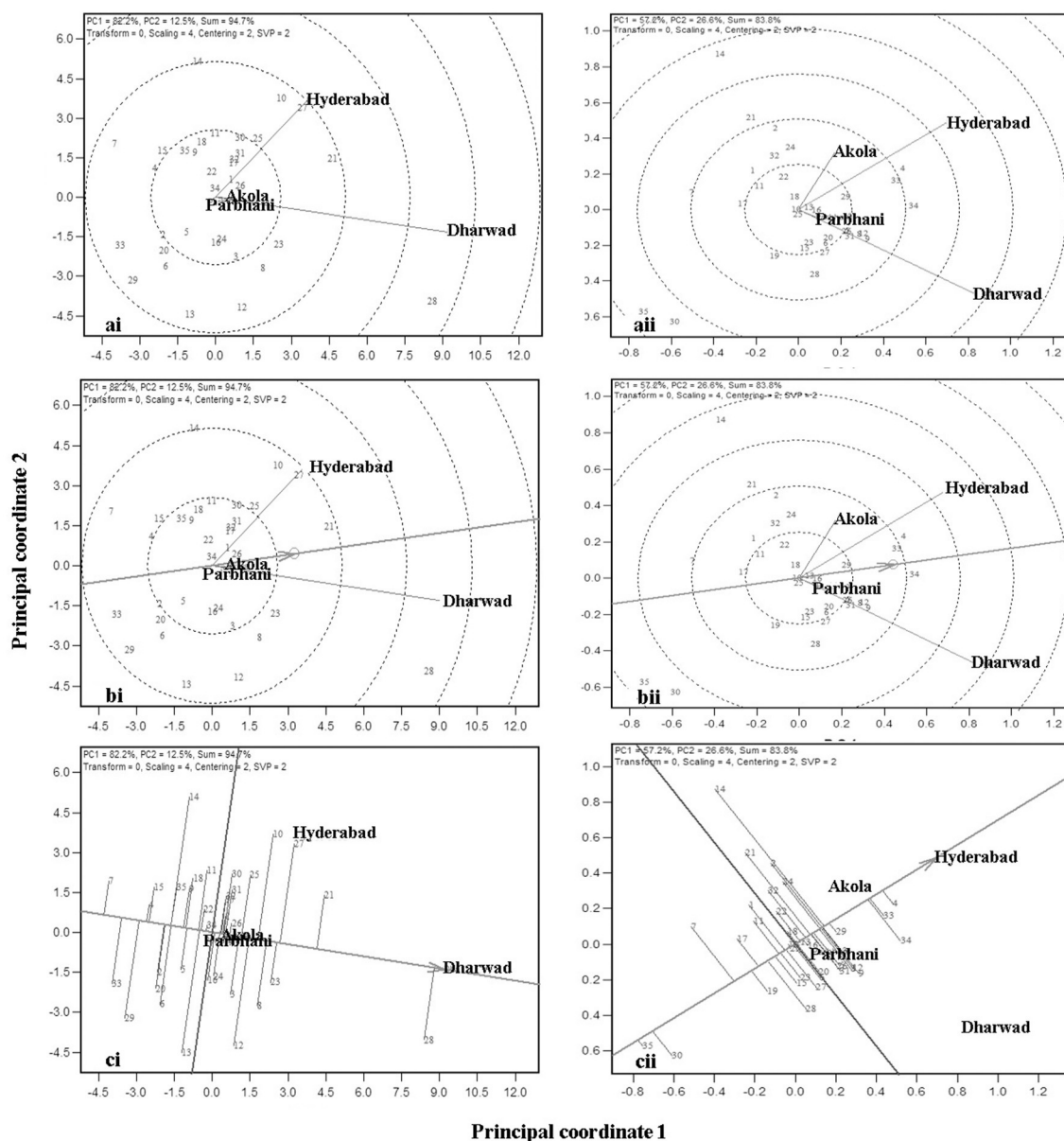


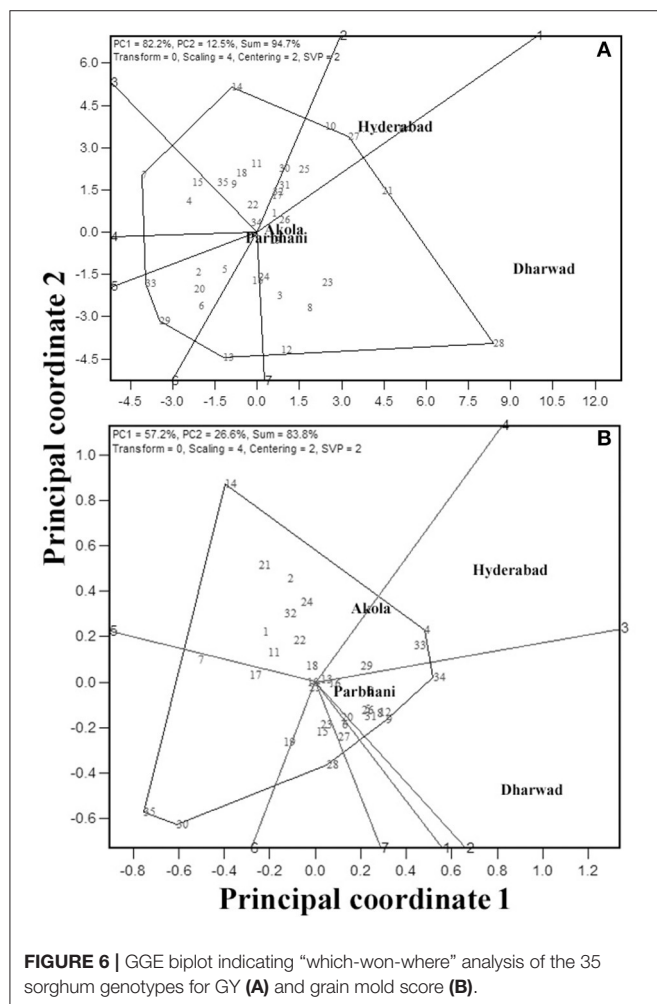
FIGURE 5 | GGE biplot showing relation among the four test locations with respect to GY (ai) and grain mold score (aii); ranking of locations based on the discriminating ability and representativeness for GY (bi) and grain mold score (bii); and Ranking of 35 sorghum genotypes based on their performance in the near ideal location, for GY (ci) and grain mold score (cii).

angle between their vectors, as explained for the environments. The shorter vector lengths indicate a minimal variation among genotypes for the trait. The $G \times T$ biplot explained 47.1% of the total variation. The vector lengths in **Figure 7** suggest that all the traits taken in this study have contributed to a variation. Based on the relationships among these traits, appropriate planning can be done in the breeding for yield and grain mold resistance. The important associations revealed by these biplots were: (i) GY has a strong positive association with flowering time, HGW, and PH,

as indicated by the acute angles. (ii) Similarly, GMS had a positive association with PH and glume coverage as indicated by the acute angles between their vectors. Thus, indicating that taller lines and lines with more GC are resistant to grain mold. The right angle between GY and GMS indicates an independent relationship among these traits. The correlation coefficients worked out with the pooled data over environments also indicated the similar results (**Table 7**), supporting the results from the GT biplot with a positive association of GY with flowering time, HGW, and

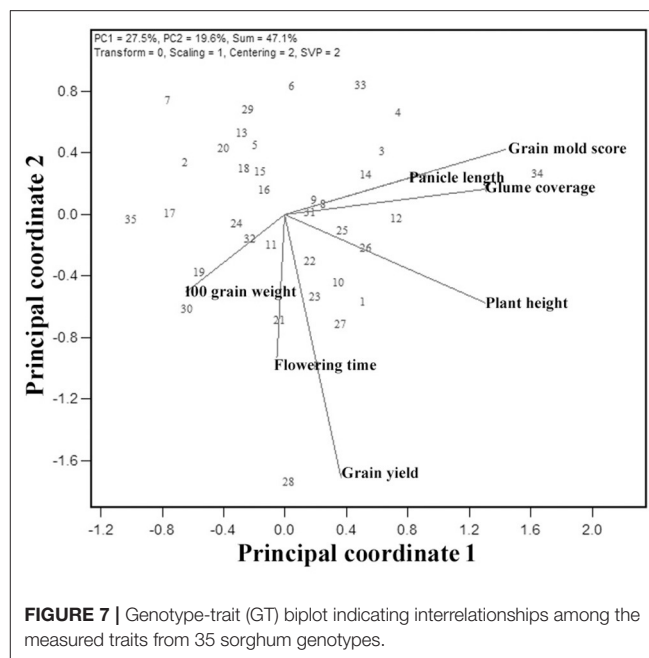
TABLE 6 | Rank correlation between environments for GY and grain mold score.

Environment	Akola	Dharwad	Hyderabad	Parbhani
Grain yield				
Akola	1			
Dharwad	0.295	1		
Hyderabad	0.180	0.572**	1	
Parbhani	-0.155	0.247	-0.382*	1
Grain mold score				
Akola	1			
Dharwad	0.123	1		
Hyderabad	0.211	0.50**	1	
Parbhani	0.038	0.432*	0.107	1

**FIGURE 6 |** GGE biplot indicating “which-won-where” analysis of the 35 sorghum genotypes for GY (A) and grain mold score (B).

PH. GMS had a significant positive correlation with PH and glume coverage, and the association of GMS was observed to be non-significant with GY.

The “Which is best for what” analysis similar to which-won-where helped to identify the superior genotypes for particular trait(s) (Figure 8). The GT biplot shows that G34 was best for grain mold resistance, glume coverage, PH, and PL, whereas G28

**FIGURE 7 |** Genotype-trait (GT) biplot indicating interrelationships among the measured traits from 35 sorghum genotypes.

was the best genotype for GY, flowering, and 100-seed weight. The genotypes that were found to have reasonably desirable grain mold score and more than average GYs were G8, G9, G31, G25, G26, and G12 from sector 1 and G10, G23, and G27 in sector 2.

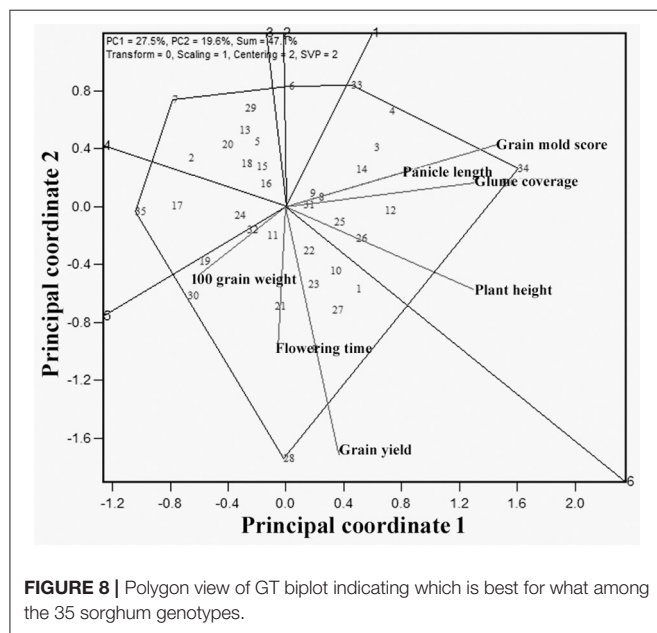
DISCUSSION

Grain mold is the production constrain of rainy season sorghum resulting in the physical and chemical deterioration of the grain, which leads to the reduction in grain size and blackening of grain, and thus making them unfit for human consumption. The molded grains cannot be used for food due to reduced processing qualities. The quality of food grains affected with grain mold is drastically decreased, and thus improving the resistance to this disease, which is a major breeding concern. The use of host-plant resistance through the resistance cultivars is the most cost-effective and eco-friendly practice among the available recommended management strategies to control mold infestation (Mofokeng et al., 2017). Global research on sorghum grain mold emphasized on the identification of disease resistance-associated grain properties, such as GHR due to a more corneous percentage of endosperm, loose panicle, increased glume coverage, pigmented glume and testa, red pericarp, phenolic compounds, and flavan-4-ols (Esele et al., 1993, 1995; Bandyopadhyay et al., 2000; Aruna and Audilakshmi, 2004; Audilakshmi et al., 2005; Thakur et al., 2006; Sharma et al., 2010; Little et al., 2012; Thirumala Rao et al., 2012). Thus, the mechanisms governing grain mold resistance are quite numerous with each contributing meagerly and making the task of resistance breeding more difficult. Genotypic variations are available for each of these traits, but combining the desirable levels of these traits into one genotype is practically challenging.

TABLE 7 | Coefficient of correlation between the studied traits across environments during 2018.

Trait	DF	PH	GC	PL	GY	HGW	GMS
DF	1						
PH	−0.072	1					
GC	0.002	0.206	1				
PL	0.077	0.174	0.272	1			
GY	0.35*	0.312	0.028	−0.06	1		
HGW	−0.296	−0.02	−0.14	0.05	0.23	1	
GMS	−0.12	0.426**	0.430**	0.069	−0.042	−0.253	1

* $p < 0.05$; ** $p < 0.01$; DF, Days to flower; PH, Plant height; GC, Glume cover; PL, Panicle length; GY, Grain yield; HGW, 100-grain weight; GMS, Grain mold score.

**FIGURE 8** | Polygon view of GT biplot indicating which is best for what among the 35 sorghum genotypes.

Successful resistance breeding program should be able to develop stable resistant lines with higher yields. Many breeding efforts were made to conquer the problem of grain mold by the pedigree breeding method using *zerazera/caudatum* line and many other grain mold-resistant sources (Ambekar et al., 2011; Prom et al., 2020). In the present study, 33 derivatives from the population breeding program are evaluated in the four testing locations. The differential ranking of genotypes across locations (hereafter referred to as environments) is caused due to a crossover GEI (Mohammadi and Amri, 2013), and a significant GEI for yield and grain mold in sorghum has been reported earlier (Rakshit et al., 2012; Aruna et al., 2016; Diatta et al., 2019).

Performance of Genotypes and Heritability

For GY, the location component explained 51.7% of the total variation, while for grain mold score, the $G \times L$ component contributed to a higher proportion of variation. The greater influence of environment on sorghum phenology and grain traits, including grain mold, was reported earlier (Kenga et al., 2006; Gasura et al., 2015; Diatta et al., 2019). The h^2 bs of all the traits

was generally observed to be high (≥ 0.50) but there was a slight difference in the individual location. Moderate to high h^2 b for DF and PH (Kenga et al., 2006; Jimmy et al., 2017) and for grain traits in different studies were reported (Almeida Filho et al., 2014; Mohammed et al., 2015; Belay and Meresa, 2017; Phuke et al., 2017). For grain mold resistance, Rodriguez-Herrera et al. (2007) and Diatta et al. (2019) reported high h^2 b, whereas Audilakshmi et al. (2011) reported low h^2 b (0.24–0.26). The high h^2 b found for most of the traits in this study indicates the additive gene action controlling these traits, which can be improved through selection (Kamatar et al., 2015).

GGE Biplot Analysis for GY and Grain Mold Resistance

The GGE and GT biplots are useful for identifying ideal genotypes and test locations, deriving the relationships among traits, and detecting the best genotypes for different traits. In this study, the complexity of the GEI for GY and grain mold resistance was investigated using the GGE biplot analysis. Highly significant mean sum of squares (Table 2) indicated the significant variability among the genotypes, the prominence of variation due to locations, and a differential response of the genotypes to the environments for GY, grain mold resistance. More than 70% of the variability for most of the traits was explained by the first two PCs. The $G + GL$ explained $>10\%$ of variability for GY and grain mold score indicates that the biplots can be used for further interpretation. For GY, the magnitude of genotype effects and GEI tested across four locations were smaller than that of environment main effect. It was observed that both crossover and non-crossover types of GEI were exhibited by the genotypes.

Information on genotype performance is also provided by the GGE biplot. Among the genotypes evaluated, G28 was the highest-yielding genotype followed by G21 (Figure 2). However, G28 was less stable when compared with other genotypes, whereas G21 expressed good stability across locations for GY. The performance of G28 perhaps suggest that a high-yielding genotype not necessarily a broad adaptation, hence it is suitable for specific adaptation. Otherwise, most stable ones are might not be high yielding (Rakshit et al., 2012; Bose et al., 2013; Aruna et al., 2016). For grain mold resistance, besides the resistant source B 58586 (G34), G4, and G33 were found to perform better and they also showed high stability for resistance.

The identification of an ideal genotype (one with high performance and greater stability) in the graphical presentation of GEI is an important application of the biplot (Yan and Tinker, 2006). From the graph, the genotype with the longest vector length and zero GEI, and located at the center of the concentric circles as in **Figure 3** is taken as an ideal genotype. Genotypes located closer to the “ideal genotype” are more desirable than those placed far away. Thus, G28, G21, G27, and G10 were closer to an ideal genotype for GY, whereas G34, G4, G33 for grain mold resistance. It is difficult to visualize from the mean table. These genotypes can be used in the breeding program and further improved for the development of stable high GY lines with better pyramiding of grain mold resistance.

The highest-yielding genotype, G28 performed the best at Dharwad, above average performance at Hyderabad, and average performance at Akola and Parbhani. The best resistant line, G34 has given its best performance at Dharwad, and in all other places its performance was good and above average. The closer angles between the trait vectors represented a closer positive relationship. The study showed that Hyderabad, Dharwad, and Akola were positively correlated, whereas Parbhani had a negative association with the other three locations for GY indicating the existence of crossover GEI for GY. The ranking of genotypes has not changed much between Hyderabad, Dharwad, and Akola, but it changed in the Parbhani location. The genotypes showing a mixture of crossover and non-crossover types of GEI in MET data is commonly reported (Fan et al., 2007; Sabaghnia et al., 2008; Rao et al., 2011; Rakshit et al., 2012; Aruna et al., 2016) which is attributed to the genotypic response to varying environments such as cropping season, rainfall, temperature regime, RH, and light intensity apart from the geographical situation of locations (latitude, longitude, and altitude) (Saeed and Francis, 1984; Dehghani et al., 2006). For grain mold resistance, Dharwad, Parbhani, and Hyderabad were positively correlated, whereas Akola had no linear correlation with Dharwad and Parbhani. In both Dharwad and Hyderabad, the rainfall and humidity were higher during the harvest months of September and October in both the years 2017 and 2018 (**Figure 1**). Hence, the grain mold incidence was high at these two locations, which are the ideal locations for carrying out grain mold screening. This result implies that the resistant entry identified in Hyderabad may reflect similar *per se* performance in Dharwad and Parbhani and *vice versa*. These locations show a good discrimination among genotypes as well as represent other testing locations and serve as appropriate breeding and testing locations. The ideal locations are those that show a good discrimination among genotypes as well as represent other testing locations. These are the locations with good general adaptation and may reduce the cost of experimentation and also improves selection efficiency.

The “which-won-where” biplot graphically summarizes the crossover GEI through creating different environmental clusters and identifying the genotypes with specific adaptation (Gauch and Zobel, 1997; Yan et al., 2000; Yan and Tinker, 2006; Putto et al., 2008; Rao et al., 2011; Rakshit et al., 2012; Aruna et al.,

2016). A genotype suitable for one or more locations can be recommended for cultivation. Based on this analysis, the present testing locations were partitioned into three clusters for GY. Cluster 1 holds Dharwad and Akola with G28 as the best genotype. Cluster 2 is Hyderabad with G27 as the best genotype, and Cluster 3 is Parbhani with G13 as the best genotype. For grain mold score, the test locations were partitioned into two clusters. Cluster 1 had Dharwad and Parbhani with G34 as the best genotype and Cluster 2 had Hyderabad and Akola with G4 as the best genotype. Thus, partitioning the target environments into different clusters and placing different genotypes specific for different clusters is the best way to exploit the positive GEI (Yan and Tinker, 2005). Such studies are available in sorghum (Rao et al., 2011; Rakshit et al., 2012). The identification of such location clusters has an advantage in region-specific grain sorghum breeding selections and cultivation.

The information generated through the GT biplot can be used to aid genotype selection based on the specific traits and helps to decide on the trait index to be assigned to these traits in the selection process. A key indicative of the GT biplot is the interrelationships among different traits and a comparison of genotypes based on the multiple traits. The GT polygon also indicated that G34 is promising for grain mold resistance, glume coverage, PH, and PL, whereas G28 is promising for GY, flowering, and 100-seed weight. G8, G9, G31, G25, G26, G12, G23, and G27 were found to have a better grain mold tolerance and more than average GYs (**Table 8**). Most of the promising genotypes with low grain mold score were found to have colored glume, supporting the earlier studies (Nida et al., 2019). As the glume covers the grains, it may shield the grain from fungal invasion, which are later on supported by the chemical and physical properties of grain thus imparting resistance to grain mold (Stenhouse et al., 1997). GCL showed a strong association with mold resistance (Ghorade et al., 2014). Generally, darker glume provides better mold resistance. A negative correlation of grain mold incidence with GC ($r = -0.56$) was reported. It may be possible to enhance grain mold resistance in white-grained sorghum by incorporating a colored glume character. Panicle structure (loose, semi-compact, and compact) determines the microclimate around the seed during the postinfection of mold colonization. Loose panicle generally dries quicker after rain than a compact one and thus influence mold development. For instance, grain mold-resistant sources are usually having loose panicles. In this study, some genotypes were identified with resistance at par with the resistant sources besides the acceptable agronomic traits like semi-compact panicle, bold seed, and less GC, indicating that the population breeding approach helps in the accumulation of favorable alleles responsible for desirable traits into a designer genotype. This approach can also be used for essentially derived varieties for grain mold and GY traits.

The $G \times T$ plot indicated that GY and GMS are genetically independent of each other suggesting that an improvement for both GY and grain mold resistance can be achieved together. This is also in conformity with the idea that current practices of cultural/chemical disease control measures can

TABLE 8 | Some lines promising for both GY and grain mold.

S. No	Genotype	Grain yield	Grain mold score	DF	PH	100 grain wt	Grain hardness	Glume cover	Glume color	Panicle compactness
1	G28	71.58	4.17	80	191	2.69	7.17	51	Light red	SC
2	G21	61.29	4.57	76	152	2.36	9.33	70	White	SL
3	G23	55.63	3.83	77	174	2.36	7.33	57	Light brown	SL
4	G27	53.83	3.56	79	198	2.36	9.17	56	Light brown	SC
5	G8	52.35	3.40	70	166	2.59	9.83	50	White	SC
6	G25	49.88	3.61	77	180	3.61	9.17	45	Red	SC
7	G31	48.52	3.31	75	150	2.39	8.33	51	Brown	SL
8	G26	47.77	3.74	73	221	2.41	7.33	50	Red	SL
9	G12	46.96	3.47	73	223	2.24	9.67	44	Brown	Loose
10	G9	42.5	3.43	77	170	2.38	9.17	49	Red	SC
	B 58586	38.88	2.52	79	202	1.78	9.67	90	Light red	Loose
	296 B	42.75	7.75	79	134	2.35	7.83	36	White	SC
	Isd	9.84	0.90	3.1	8.08	0.24		7.45	-	-

SC, Semi-compact; SL, Semi loose.

improve the yield by controlling grain mold. Based on the trait associations, GY improvement could be achieved through PH, flowering time, and HGW using a correlated response to selection. Plant and GY component traits such as grain weight and 1,000-grain weight contributed toward higher GY (Tovignan et al., 2016; da Silva et al., 2017). Though a few earlier studies reported a negative correlation between grain mold resistance and yield components (Reddy et al., 2008; Diatta et al., 2019), they could find some genotypes that could combine good GY and grain mold tolerance, suggesting the possibility to break the negative linkage by growing and exercising selection in large segregating populations (Ambekar et al., 2011; Diatta et al., 2019). A previous study reported that climatic factors like high RH (85–100%) and temperature (25–30°C) prevail during crop maturity (September) favors the natural mold infestation (Tonapi et al., 2007). In such conditions, if there are some genotypes having low infestation, this could be attributed to their level of tolerance to mold infestation in comparison to other genotypes. Hence, in addition to genetic parameters, weather parameters in testing season are mandatory in grain mold resistance breeding nurseries and trials.

CONCLUSIONS

The present study assessed the multilocation field performance of 33 derivatives of the population breeding approach, for their GY and grain mold resistance. The results showed a significant GEI on both GY and grain mold. The h^2 bs was generally high, indicating the possibility of improvement through a progeny selection for these traits. The GGE biplot analysis helped to identify G28 and G21 as the best candidates for GY, whereas G4, G33 were good for grain mold resistance in terms of their mean performance and stability compared to other lines. The study also identified the lines with high GY

and moderately high grain mold resistance. The study further suggested that Dharwad is an ideal location for the evaluation of both GY and grain mold resistance. It is proved that population breeding, by involving diverse and elite parents, is effective in bringing together agronomically desirable characters thus help in improving grain quality in terms of grain mold resistance. Therefore, the demand for rainy season sorghum will increase for food, feed, and starch industries. Genetic material in this study also merits further genomic studies and marker development for mold tolerance. Hence, any improvement in grain quality cannot be separated from the improvement for grain mold resistance, which benefits the farmers with premium price in the markets.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

CA: conceptualization, data curation, writing—original draft, and manuscript preparation. ID: execution of the experiment at different centers and manuscript preparation. PR: analysis of the data, preparation of tables, graphs, and manuscript preparation. RG and VK: execution of the experiment at Akola and the collection of field data. AG: execution of the experiment at Akola and scoring of samples for grain mold resistance. SK and NH: execution of the experiment at Dharwad and the collection of field data. SC: execution of the experiment at Dharwad and scoring of samples for grain mold resistance. SM and KK: execution of the experiment at Parbhani and the collection of field data. VG: execution of the experiment at Parbhani and scoring of samples for

grain mold resistance. CD and DB: field experimentation and data collection at Hyderabad. NK: recording data on GHR at Hyderabad. MG: data interpretation and manuscript preparation. VT: project administration and resources. All authors contributed to the article and approved the submitted version.

REFERENCES

- Almeida Filho, J. E. D., Tardin, F. D., Daher, R. F., Barbe, T. C., Paula, C. M., Cardoso, M. J., et al. (2014). Stability and adaptability of grain sorghum hybrids in the off-season. *Genet. Mol. Res.* 13, 7626–7635. doi: 10.4238/2014.March.24.24
- Ambekar, S. S., Kamatar, M. Y., Ganesamurthy, K., Ghorade, R. B., Saxena, U., Chand, P., et al. (2011). Genetic enhancement of Sorghum (*Sorghum bicolor* (L.) Moench) for grain mould resistance: II. breeding for grain mould resistance. *Crop Protect.* 30, 759–764. doi: 10.1016/j.cropro.2010.06.024
- Aruna, C., and Audilakshmi, S. (2004). Genetic architecture of grain hardness – a durable resistance mechanism for grain moulds in sorghum [*Sorghum bicolor* (L.) Moench]. *Indian J. Genet. Plant Breed.* 64, 35–38.
- Aruna, C., Rakshit, S., Shrotria, P. K., Shiva Kumar, S. K. P., Modi, N. D., Deshmukh, D. T., et al. (2016). Assessing genotype-by-environment interactions and trait associations in forage sorghum using GGE biplot analysis. *J. Agric. Sci.* 154, 73–86. doi: 10.1017/S0021859615000106
- Aruna, C., Visarada, K. B. R. S., Bhat, B. V., and Tonapi, V. A. (eds.). (2019). *Breeding Sorghum for Diverse end Uses, 1st Edn.* Woodhead Publishing, 434. doi: 10.1016/B978-0-08-101879-8.00001-2
- Ashok Kumar, A., Reddy, B. V. S., Sharma, H. C., Hash, C. T., Rao, P. S., Ramaiah, B., et al. (2011). Recent advances in sorghum genetic enhancement research at ICRISAT. *Am. J. Sci.* 2, 589–600. doi: 10.4236/ajps.2011.24070
- Audilakshmi, S., Das, I. K., Ghorade, R. B., Mane, P. N., Kamatar, M. Y., Narayana, Y. D., et al. (2011). Genetic improvement of sorghum for grain mould resistance: I. performance of sorghum recombinant inbred lines for grain mould reactions across environments. *Crop Protect.* 30, 753–758. doi: 10.1016/j.cropro.2010.12.024
- Audilakshmi, S., Stenhouse, J. W., and Reddy, T. P. (2005). Genetic analysis of grain mold resistance in white seed sorghum genotypes. *Euphytica* 145, 95–101. doi: 10.1007/s10681-005-0534-6
- Bandyopadhyay, R., Butler, D. R., Chandrashekar, A., Kanaka, R., and Navi, S. S. (2000). “Biology, epidemiology, and management of sorghum grain mold,” in *Technical and Institutional Options for Sorghum Grain Mold Management*, eds A. Chandrashekar, R. Bandyopadhyay, and A. J. Hall (Patancheru: ICRISAT), 34–51.
- Belay, F., and Meresa, H. (2017). Performance evaluation of sorghum [*Sorghum bicolor* (L.) Moench] hybrids in the moisture stress conditions of Abergelle District, Northern Ethiopia. *J. Cereals Oilseeds* 8, 26–32. doi: 10.5897/CO2016.0168
- Bernardino, K. C., de Menezes, C. B., de Sousa, S. M., Guimarães, C. T., Carneiro, P. C. S., Schafert, R. E., et al. (2021). Association mapping and genomic selection for sorghum adaptation to tropical soils of Brazil in a sorghum multiparental random mating population. *Theor. Appl. Genet.* 134, 295–312. doi: 10.1007/s00122-020-03697-8
- Bhola, N. (1982). “Population breeding techniques in sorghum,” in *Sorghum in eighties. Proceedings of the International Symposium on Sorghum, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), 2-7 November 1981*, Patancheru, AP, India, eds L. R. House, L. K. Mughogho, and J. M. Peacock (Patancheru: ICRISAT), 421–434.
- Bose, L. K., Jambhulkar, N. N., Pande, K., and Singh, O. N. (2013). Use of AMMI and other stability statistics in the simultaneous selection of rice genotypes for yield and stability under direct-seeded conditions. *Chil. J. Agric. Res.* 74. doi: 10.4067/S0718-58392014000100001
- Castor, L. L., and Frederiksen, R. A. (1980). “Fusarium and Curvularia grain mold in Texas,” in *Sorghum Diseases, A World Review. Proceedings of an International Workshop on Sorghum Diseases, 11–15 December, 1978*, eds R. J. Williams, R. A. Frederiksen, L. K. Mughogho, and G. D. Bengston (Hyderabad: ICRISAT), 93–102
- Comstock, R. E., and Robinson, H. F. (1952). “Genetic parameters, their estimation and significance,” in *Proceedings of the Sixth International Grassland Congress* (State College, PA: Pennsylvania State College), 284–291.
- Cooper, M., Delacy, I. H., and Basford, K. E. (1996). “Relationships among analytical methods used to analyse genotypic adaptation in multi-environment trials,” in *Plant Adaptation and Crop Improvement*, eds M. Cooper and G. L. Hammer (Wallingford: CAB International), 193–224.
- da Silva, K. J., Teodoro, P. E., de Menezes, C. B., Júlio, M. P. M., de Souza, V. F., da Silva, M. J., et al. (2017). Contribution of morpho agronomic traits to grain yield and earliness in grain sorghum. *Genet. Mol. Res.* 16, 1–10. doi: 10.4238/gmr16029649
- Das, I. K. (2019). “Advances in sorghum disease resistance,” in *Breeding Sorghum for Diverse end Uses*, eds C. Aruna, K. B. R. S. Visarada, V. B. Bhat and V. A. Tonapi (Woodhead Publishing Series in Food Science, Technology and Nutrition), 313–324, 434.
- Das, I. K., Aruna, C., and Tonapi, V. A. (2020). *Sorghum Grain Mold*. Hyderabad: ICAR-Indian Institute of Millets Research, 86.
- Das, I. K., Audilakshmi, S., and Patil, J. V. (2012). Fusarium grain mold: the major component of grain mold disease complex in sorghum (*sorghum bicolor* L. Moench). *Eur. J. Plant Sci. Biotechnol.* 6, 45–55.
- Das, I. K., and Padmaja, P. G. (2016). *Biotic Stress Resistance in Millets*. Academic Press, 246.
- Das, I. K., and Patil, J. V. (2013). “Assessment of economic loss due to grain mold of sorghum in India,” in *Compendium of Paper and Abstracts: Global Consultation on Millets Promotion for Health and Nutritional Security, 18-20 Dec., 2013*, eds S. Rakshit, I. K. Das, G. Shyamprasad, J. S. Mishra, C. V. Ratnavathi, R. R. Chapke (Hyderabad: Society for Millet Research, Directorate of Sorghum Research), 59–63.
- Dehghani, H., Ebadi, A., and Yousefi, A. (2006). Biplot analysis of genotype by environment interaction for barley yield in Iran. *Agron. J.* 98, 388–393. doi: 10.2134/agronj2004.0310
- Dehghani, H., Omidi, H., and Sabaghnia, N. (2008). Graphic analysis of trait relations of rapeseed using the biplot method. *Agron. J.* 100, 1443–1449. doi: 10.2134/agronj2007.0275
- Diatla, C., Tovignan, T. K., Adoukonou-Sagbadja, H., Aidara, O., Diao, Y., Penda Sarr, M., et al. (2019). Development of sorghum hybrids for stable yield and resistance to grain mold for the Center and South-East of Senegal. *Crop Protect.* 119: 197–207. doi: 10.1016/j.cropro.2019.02.001
- Doggett, H. (1972). Recurrent selection in sorghum populations. *Heredity* 28, 9–29. doi: 10.1038/hdy.1972.2
- Eberhart, S. A. (1972). “Techniques and methods of more efficient population improvement in sorghum. Pages 198–213,” in *Sorghum in Seventies*, eds N. G. P. Rao, and L. R. House (New Delhi: Oxford and IBH Publishing Co), 198–213.
- Esele, J. P., Frederiksen, R. A., and Miller, F. R. (1993). The association of genes controlling caryopsis traits with grain mould resistance in sorghum. *Phytopathology* 83, 490–495. doi: 10.1094/Phyto-83-490
- Esele, J. P., Frederiksen, R. A., and Miller, F. R. (1995). The importance of plant color and modifier genes in grain mold resistance in sorghum. *East African Agric. For. J.* 61, 31–37. doi: 10.4314/eaaf.v61i1.46789
- Fan, X. M., Kang, M. S., Chen, H., Zhang, Y., Tan, J., and Xu, C. (2007). Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. *Agron. J.* 99, 220–228. doi: 10.2134/agronj2006.0144
- Forbes, G. A., Bandyopadhyay, R., and Garcia, G. (1992). “A review of sorghum grain mold,” in *Sorghum and Millet Diseases; A Second World Review*, eds J. W. A. de Milliano, R. A. Frederiksen, and G. D. Bengston (Patancheru: ICRISAT), 253–264.
- Frederiksen, R. A., Castor, L. L., and Rosenow, D. T. (1982). “Grain mold, small seed and head blight: the Fusarium connection in sorghum,” in

ACKNOWLEDGMENTS

We sincerely acknowledge Indian Council of Agricultural Research (ICAR), New Delhi and ICAR-IIMR for their physical and financial support to this study, and ICRISAT for its research and development support.

- Proceedings of the Thirty Seventh Annual Corn and Sorghum Industry Research Conference*, 26–36.
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58, 453–467. doi: 10.1093/biomet/58.3.453
- Gasura, E., Setimela, P. S., and Souta, C. M. (2015). Evaluation of the performance of sorghum genotypes using GGE biplot. *Can. J. Plant Sci.* 95, 1205–1214. doi: 10.4141/cjps-2015-119
- Gauch, H. G., and Zobel, R. W. (1997). Identifying megaenvironment and targeting genotypes. *Crop Sci.* 37, 381–385. doi: 10.2135/cropsci1997.0011183X003700020002x
- Ghorade, R. B., Kalpande, V. V., Bhongle, S. A., and Bhakre, R. L. (2014). Interrelationship among the parameters associated with grain mold resistance in kharif sorghum. *Plant Arch.* 14, 487–489.
- Glueck, J. A., and Rooney, L. W. (1980). “Chemistry and structure of grain in relation to mold resistance,” in *Sorghum Diseases: A World Review. Proceedings of an International Workshop at ICRISAT*, eds J. R. Williams, R. A. Frederiksen, L. K. Mughogho, and G. D. Bengston (Patancheru), 119–140.
- Gonzalez, R., Phillips, R., Saloni, D., Jameel, H., Abt, R., Pirraglia, A., et al. (2011). Biomass to energy in the Southern United States: supply chain and delivered cost. *BioResources* 6, 2954–2976.
- ICRISAT (1992). *ICRISAT Medium Term Plan 1994-98 Research Theme Datasets*. Patancheru: ICRISAT, 229.
- Indira, S., Muthusubramanian, V., Tonapi, V. A., Varanavasiappan, S., and Seetharama, N. (2006). Utility of biocontrol agents in the suppression of seed borne pathogenic mycoflora and their effect on seed quality in sorghum (*Sorghum bicolor* (L.) Moench). *Arch. Phytopathol. Plant Prot.* 39, 379–387. doi: 10.1080/03235400500301158
- Jimmy, M., Nzuve, F., Florence, O., Manyasa, E., and Muthomi, J. (2017). Genetic variability, heritability, genetic advance and trait correlations in selected sorghum (*Sorghum bicolor* L. Moench) varieties. *Int. J. Agron. Agric. Res.* 11, 47–56.
- Kamatar, M., Biradar, B., Sowmya, H., Brunda, S., Deepakkumar, G. S., Rajaput, S., et al. (2015). Genetic variability and selection for grain mold resistance in F3 progenies of sorghum. *Int. J. Farming Allied Sci.* 4, 334–340.
- Kenga, R., Tenkouano, A., Gupta, S. C., and Alabi, S. O. (2006). Genetic and phenotypic association between yield components in hybrid sorghum (*Sorghum bicolor* (L.) Moench) populations. *Euphytica* 150, 319–326. doi: 10.1007/s10681-006-9108-5
- Klein, R. R., Rodriguez-Herrera, R., Schlueter, J., Klein, P. E., Yu, Z. H., and Rooney, W. L. (2001). Identification of genomic regions that affect grain-mould incidence and other traits of agronomic importance in sorghum. *Theor. Appl. Genet.* 102, 307–319. doi: 10.1007/s001220051647
- Little, C. R., and Magill, C. W. (2009). The grain mold pathogen, *Fusarium thapsinum*, reduces premature seed rot caused by Fusarium grain mold in sorghum. *Indian J. Plant Prot.* 41, 238–243. doi: 10.1111/j.1439-0434.2008.01530.x
- Little, C. R., Perumal, R., Tesso, T., Prom, L. K., Odvody, G. N., and Mangill, C. W. (2012). Sorghum pathology and biotechnology - a fungal disease perspective: Part I. Grain mold, head smut and ergot. *Eur. J. Plant Sci. Biotechnol.* 6, 10–30.
- Malla, S., Ibrahim, A. M. H., Little, R., Kalsbeck, S., Glover, K. D., and Ren, C. (2010). Comparison of shifted multiplicative model, rank correlation, and biplot analysis for clustering winter wheat production environments. *Euphytica* 174, 357–370. doi: 10.1007/s10681-010-0130-2
- Maunder, A. B. (1972). “Objectives and approaches to grain and forage sorghum improvement in the Americas. Pages 60–100,” in *Sorghum in Seventies*, eds N. G. P. Rao, and L. R. House (New Delhi: Oxford and IBH Publishing Co. Pvt Ltd), 60–100.
- Mofokeng, M. A., Shimelis, H., Laing, M., and Shargie, N. (2017). Sorghum [*Sorghum bicolor* (L.) Moench] breeding for resistance to leaf and stalk anthracnose, *Colletotrichum sublineolum*, and improved yield: progress and prospects. *Aust. J. Crop. Sci.* 11, 1078–1085. doi: 10.21475/ajcs.17.11.09.pne347
- Mohammadi, R., and Amri, A. (2013). Genotype × environment interaction and genetic improvement for yield and yield stability of rainfed durum wheat in Iran. *Euphytica* 192, 227–249. doi: 10.1007/s10681-012-0839-1
- Mohammed, R., Ashok Kumar, A., Bhavanasi, R., Munghate, R. S., Kavi Kishor, P. B., and Sharma, H. C. (2015). Quantitative genetic analysis of agronomic and morphological traits in sorghum, *Sorghum bicolor*. *Front. Plant Sci.* 6:945. doi: 10.3389/fpls.2015.00945
- Mpofu, L. T., and McLaren, N. W. (2014). Ergosterol concentration and variability in genotype-by-pathogen interaction for grain mold resistance in sorghum. *Planta* 240: 239–250. doi: 10.1007/s00425-014-2081-7
- Murty, U. R., and Rao, N. G. P. (1997). “Sorghum,” in *Genetics, Cytogenetics and Breeding of Crop plants (Vol. 2), Cereal and Commercial Crops*, eds P. N. Bahl, P. M. Salimath, and A. K. Mandal (New Delhi: Oxford & IBH Publishing Co. Pvt. Ltd), 198–239.
- Navi, S. S., Bandyopadhyay, R., and Hall, A. J. (2002). Effects of crop season, storage conditions, cultivars, and fungicide on postharvest mold fungi infecting sorghum grain. *ISMN* 43, 65–68.
- Navi, S. S., Bandyopadhyay, R., Reddy, R. K., Thakur, R. P., and Yang, X. B. (2005). Effects of wetness duration and grain development stages on sorghum grain mold infection. *Plant Dis.* 89, 872–878. doi: 10.1094/PD-89-872
- Nida, H., Girma, G., Mekonen, M., Lee, S., Seyoum, A., Dessalegn, K., et al. (2019). Identification of sorghum grain mold resistance loci through genome wide association mapping. *J. Cereal Sci.* 85, 295–304. doi: 10.1016/j.jcs.2018.12.016
- Phuke, R. M., Anuradha, K., Radhika, K., Jabeen, F., Anuradha, G., Ramesh, T., et al. (2017). Genetic variability, genotype - environment interaction, correlation, and GGE biplot analysis for grain iron and zinc concentration and other agronomic traits in RIL population of sorghum (*sorghum bicolor* L. Moench). *Front. Plant Sci.* 8:712. doi: 10.3389/fpls.2017.00712
- Prom, L. K., Cuevas, H. E., Ahn, E., Isaikit, T., Rooney, W. L., and Magill, C. (2020). Genome wide association study of grain mold resistance in sorghum association panel as affected by inoculation with *Alternaria alternata* alone and *Alternaria alternata*, *Fusarium thapsinum*, and *Curvularia lunata* combined. *Eur. J. Plant Pathol.* 157, 783–798. doi: 10.1007/s10658-020-02036-3
- Prom, L. K., and Erpelding, J. E. (2009). New sources of grain mold resistance among sorghum accessions from Sudan. *Trop. Subtropical Agroecosyst.* 10, 457–463.
- Prom, L. K., Perumal, R., Cisse, N., and Little, C. R. (2014). Evaluation of selected sorghum lines and hybrids for resistance to grain mold and long smut fungi in Senegal, West Africa. *Plant Health Prog.* 15, 28–31. doi: 10.1094/PHP-RS-13-0128
- Prom, L. K., Waniska, R. D., Kollo, A. I., and Rooney, W. L. (2003). Response of eight sorghum cultivars inoculated with *Fusarium thapsinum*, *curvularia lunata* and a mixture of the two fungi. *Crop Protect.* 22, 623–628. doi: 10.1016/S0261-2194(02)00246-6
- Putto, W., Patanothai, A., Jogloy, S., and Hoogenboom, G. (2008). Determination of mega-environments for peanut breeding using the CSM-CROPGRO-Peanut model. *Crop Sci.* 48, 973–982. doi: 10.2135/cropsci2007.1.0.0552
- Rakshit, S., Ganapathy, K. N., Gomashe, S. S., Rathore, A., Ghorade, R. B., Kumar, M. V. N., et al. (2012). GGE biplot analysis to evaluate genotype, environment and their interactions in sorghum multi-location data. *Euphytica* 185, 465–479. doi: 10.1007/s10681-012-0648-6
- Rao, P. S., Reddy, P. S., Rathore, A., Reddy, B. V. S., and Panwar, S. (2011). Application of GGE biplot and AMMI model to evaluate sweet sorghum (*Sorghum bicolor*) hybrids for genotype-environment interaction and seasonal adaptation. *Indian J. Agric. Sci.* 81, 438–444.
- Rao, V. T., Reddy, P. S., Thakur, R. P., and Reddy, B. V. S. (2013). Physical kernel properties associated with grain mold resistance in sorghum [*Sorghum bicolor* (L.) Moench]. *Int. J. Plant Breeding Genet.* 7, 176–181. doi: 10.3923/ijpb.2013.176.181
- Ratnadassa, A., Butler, D. R., Marley, P. S., Bandyopadhyay, R., Hesse, D. E., and Akintayo, I. (2003). Sorghum head- bugs and grain molds in West and Central Africa: II. relationships between weather, head bug and mold damage on sorghum grains. *Crop Protect.* 22, 853–858. doi: 10.1016/S0261-2194(03)00067-X
- Reddy, B. V. S., Ramesh, S., Kumar, A. A., Wani, S. P., Ortiz, R., Ceballos, H., et al. (2008). Bio-fuel crops Research for energy security and rural development in developing countries. *Bioenerg. Res.* 1, 248–258. doi: 10.1007/s12155-008-9022-x
- Reddy, B. V. S., Ramesh, S., Madhududhana, R., Aruna, C., and Sanjana Reddy, P. (2006). *Population improvement in sorghum. Information Bulletin no. 73*. Patancheru: International Crops Research Institute for the Semi-Arid Tropics, 64. ISBN 92-9066-492-4. Order code IBE 073.

- Reddy, B. V. S., and Stenhouse, J. W. (1994). Sorghum improvement for the semi-arid tropic region: past, current and future research thrusts in Asia. *Punjabrao Krishi Vidyapeeth Res. J.* 18, 155–170.
- Robinson, H. F., Comstock, R. E., and Harvey, P. H. (1949). Genotypic and phenotypic correlation in corn and their implications in selection. *Agron. J.* 43, 282–287. doi: 10.2134/agronj1951.00021962004300060007x
- Rodríguez, F., Alvarado, G., Pacheco, Á., and Burgueño, J. (2018) *ACBD-R Augmented Complete Block Design with R for Windows, Version 4.0 [Internet]*. CIMMYT Research Data and Software Repository, Network.
- Rodríguez-Herrera, R., Rooney, W. L., Rosenow, D., and Frederiksen, R. A. (2000). Inheritance of grain mold resistance in grain sorghum without a pigmented testa. *Crop Sci.* 40, 1573–1578. doi: 10.2135/cropsci2000.4061573x
- Rodríguez-Herrera, R., Rooney, W. L., Rosenow, D. T., Aguilar-Gonzalez, C. N., and Quero-Carrillo, A. R. (2007). Genetic control and heritability of resistance to grain mold in F2:5 sorghum families with non-pigmented testa. *Plant Breed. Seed Sci.* 55, 34–43.
- Rodríguez-Herrera, R., Waniska, R. D., and Rooney, W. L. (1999). Antifungal proteins and grain mold resistance in sorghum with nonpigmented testa. *J. Agric. Food Chem.* 47, 4802–4806. doi: 10.1021/jf9903814
- Romagosa, I., and Fox, P. N. (1993). “Genotype × environment interaction and adaptation,” in *Plant Breeding: Principles and Prospects*, eds M. D. Hayward, N. O. Bosemark, and I. Romagosa (London: Chapman & Hall), 373–390. doi: 10.1007/978-94-011-1524-7_23
- Sabaghnia, N., Dehghani, H., and Sabaghpour, S. H. (2008). Graphic analysis of genotype by environment interaction for lentil yield in Iran. *Agron. J.* 100, 760–764. doi: 10.2134/agronj2006.0282
- Saeed, M., and Francis, C. A. (1984). Association of weather variable with genotype × environment interactions in grain sorghum. *Crop Sci.* 24, 13–16. doi: 10.2135/cropsci1984.0011183X002400010004x
- Sharma, R., Rao, V. P., Upadhyaya, H. D., Reddy, V. G., and Thakur, R. P. (2010). Resistance to grain mold and downy mildew in a mini-core collection of sorghum germplasm. *Plant Dis.* 94, 439–444. doi: 10.1094/PDIS-94-4-0439
- Singh, D. P., and Bandyopadhyay, R. (2000). “Grain mould,” in *Compendium of Sorghum Diseases*, 2nd Edn, eds R. A. Frederiksen and G. N. Odvody (St Paul, MN: APS Press), 38–40.
- Stenhouse, J., Bandyopadhyay, R., Singh, S. D., and Subramanian, V. (1997). “Breeding for grain mold resistance in sorghum,” in *Proceedings of the International Conference on Genetic Improvement of Sorghum and Pearl Millet* (Lubbock, Tx), 326–336.
- Tarekegn, G., McLaren, N. W., and Swart, W. J. (2006). Effects of weather variables on grain mold of sorghum in South Africa. *Plant Pathol.* 55, 238–245. doi: 10.1111/j.1365-3059.2006.01333.x
- Thakur, R. P., Reddy, B. V. S., Indira, S., Rao, V. P., Navi, S. S., Yang, X. B., et al. (2006). *Sorghum Grain Mold. Information Bulletin No. 72*. Patancheru: International Crops Research Institute for the Semi-arid Tropics, 32.
- Thirumala Rao, V., Sanjana Reddy, P., Reddy, B. V. S., and Ashok Kumar, A. (2012). Ergosterol as an indicator of mold resistance in white and colored grain sorghum hybrids. *Crop Protect.* 37, 7–12. doi: 10.1016/j.cropro.2011.12.017
- Tonapi, V. A., Mundada, R. R., Navi, S. S., Reddy, R. K., Thakur, R. P., Bandyopadhyay, R., et al. (2007). Effect of temperature and humidity regimes on grain mold sporulation and seed quality in sorghum (*Sorghum bicolor* (L.) Moench). *Arch. Phytopathol. Plant Prot.* 40, 113–127. doi: 10.1080/03235400500355626
- Tovignán, T. K., Fonceka, D., Ndoye, I., Cisse, N., and Luquet, D. (2016). The sowing date and post-flowering water status affect the sugar and grain production of photoperiodic, sweet sorghum through the regulation of sink size and leaf area dynamics. *F. Crop. Res.* 192, 67–77. doi: 10.1016/j.fcr.2016.04.015
- Ulaganathan, K. (2011). “Biotechnological approaches for management of grain mold disease of sorghum,” in *Pests and Pathogens: Management Strategies*, eds V. Dashavantha Reddy, P. Nagaraja Rao, and K. Venkateswara Rao (BS Publications), 67.
- V. S. N., International (2009). *GenStat Software for Windows. Release 12.1*. Hemel Hempstead: VSN International Ltd.
- Waniska, R. D., Venkatesha, R. T., Chandrashekar, A., Krishnaveni, S., Bejosano, F. P., Jeoung, J., et al. (2001). Antifungal proteins and other mechanisms in the control of sorghum stalk rot and grain mold. *J. Agric. Food Chem.* 49, 4732–4742. doi: 10.1021/jf010007f
- Williams, R. J., and Rao, K. N. (1981). A review of sorghum grain mold. *Trop. Pest. Manag.* 27, 200–211. doi: 10.1080/09670878109413652
- Yan, W. (2001). GGE biplot—a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agron. J.* 93, 1111–1118. doi: 10.2134/agronj2001.9351111x
- Yan, W. (2002). Singular-value partitioning in biplot analysis of multi-environment trial data. *Agron. J.* 94, 990–996.
- Yan, W., and Holland, J. B. (2010). A heritability adjusted GGE biplot for test environment evaluation. *Euphytica* 171, 355–369. doi: 10.1007/s10681-009-0030-5
- Yan, W., Hunt, L., Sheng, Q., and Szlavics, Z. (2000). Cultivar evaluation and mega environment investigation based on the GGE biplot. *Crop Sci.* 40, 597–605. doi: 10.2135/cropsci2000.403597x
- Yan, W., and Kang, M. S. (2003). *GGE Biplot Analysis: a Graphical Tool for Breeders, Geneticists, and Agronomists*. Boca Raton, FL: CRC Press. doi: 10.1201/9781420040371
- Yan, W., Kang, M. S., Ma, B. L., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs AMMI analysis of genotype by environment data. *Crop Sci.* 47, 643–653. doi: 10.2135/cropsci2006.06.0374
- Yan, W., and Rajcan, I. (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Sci.* 42, 11–20. doi: 10.2135/cropsci2002.0011
- Yan, W., and Tinker, N. A. (2005). An integrated biplot analysis system for displaying, interpreting and exploring genotype × environment interactions. *Crop Sci.* 45, 1004–1016. doi: 10.2135/cropsci2004.0076
- Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: principles and applications. *Can. J. Plant Sci.* 86, 623–645. doi: 10.4141/P05-169

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Aruna, Das, Reddy, Ghorade, Gulhane, Kalpande, Kajjidoni, Hanamaratti, Chattannavar, Mehtre, Gholve, Kamble, Deepika, Kannababu, Bahadure, Govindaraj and Tonapi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Genome-Wide Assessment of Population Structure and Genetic Diversity of the Global Finger Millet Germplasm Panel Conserved at the ICRISAT Genebank

C. Backiyalakshmi^{1,2}, Mani Vetriventhan^{2*}, Santosh Deshpande², C. Babu¹, V. Allan¹, D. Naresh², Rajeev Gupta² and Vania C. R. Azevedo^{2*}

¹ Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University (TNAU), Coimbatore, India, ² International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Salej Sood,
Indian Council of Agricultural Research
(ICAR), India
Raj K. Pasam,
AgriBio, La Trobe University, Australia
Antonio Lupini,
Mediterranea University of Reggio
Calabria, Italy

*Correspondence:

Mani Vetriventhan
M.Vetriventhan@cgiar.org
Vania C. R. Azevedo
V.Azevedo@cgiar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 08 April 2021

Accepted: 16 June 2021

Published: 20 August 2021

Citation:

Backiyalakshmi C, Vetriventhan M,
Deshpande S, Babu C, Allan V,
Naresh D, Gupta R and Azevedo VCR
(2021) Genome-Wide Assessment of
Population Structure and Genetic
Diversity of the Global Finger Millet
Germplasm Panel Conserved at the
ICRISAT Genebank.
Front. Plant Sci. 12:692463.
doi: 10.3389/fpls.2021.692463

Finger millet [*Eleusine coracana* (L.) Gaertn.] is an important climate-resilient nutrient-dense crop grown as a staple food grain in Asia and Africa. Utilizing the full potential of the crop mainly depends on an in-depth exploration of the vast diversity in its germplasm. In this study, the global finger millet germplasm diversity panel of 314 accessions was genotyped, using the DArTseq approach to assess genetic diversity and population structure. We obtained 33,884 high-quality single nucleotide polymorphism (SNP) markers on 306 accessions after filtering. Finger millet germplasm showed considerable genetic diversity, and the mean polymorphic information content, gene diversity, and Shannon Index were 0.110, 0.114, and 0.194, respectively. The average genetic distance of the entire set was 0.301 (range 0.040 – 0.450). The accessions of the race *elongata* (0.326) showed the highest average genetic distance, and the least was in the race *plana* (0.275); and higher genetic divergence was observed between *elongata* and *vulgaris* (0.320), while the least was between *compacta* and *plana* (0.281). An average, landrace accessions had higher gene diversity (0.144) and genetic distance (0.299) than the breeding lines (0.117 and 0.267, respectively). A similar average gene diversity was observed in the accessions of Asia (0.132) and Africa (0.129), but Asia had slightly higher genetic distance (0.286) than African accessions (0.276), and the distance between these two regions was 0.327. This was also confirmed by a model-based STRUCTURE analysis, genetic distance-based clustering, and principal coordinate analysis, which revealed two major populations representing Asia and Africa. Analysis of molecular variance suggests that the significant population differentiation was mainly due to within individuals between regions or between populations while races had a negligible impact on population structure. Finger millet diversity is structured based on a geographical region of origin, while the racial structure made negligible contribution to population structure. The information generated from this study can provide greater insights into the population structure and genetic diversity within and among regions and races, and an understanding of genomic-assisted finger millet improvement.

Keywords: finger millet, germplasm, genetic diversity, population structure, DArTseq, Analysis of molecular variance (AMOVA)

INTRODUCTION

Finger millet [*Eleusine coracana* (L.)] is an important nutraceutical crop. It is highly adapted to the semiarid tropics and is grown as a staple food crop in Asia and Africa. Globally, it is the sixth most important crop among cereals in terms of production, and it contributes about 12% of the total millet area (Mundada et al., 2020). Its origin dates back to 5,000 years in western Uganda and the Ethiopian highlands. In India, its cultivation can be traced to 3,000 BC in the Western Ghats; thus, India is considered a secondary center of diversity for finger millet (Hilu and DeWet, 1976; Hilu et al., 1979). The crop is highly self-pollinated and allotetraploid (AABB) with chromosome number $2n = 4x = 36$. Finger millet is gaining importance and drawing attention globally due to its grain nutrient composition, with high dietary fiber (11–20%), essential amino acids, vitamins, and micronutrients, particularly calcium (1.8–4.9 g/kg), iron (~22–65 mg/kg), zinc (~17–25 mg/kg), protein (6–11%), carbohydrates (65–75%), and other minerals; it is also gluten-free (Chethan and Malleshi, 2007; Upadhyaya et al., 2011; Shobana et al., 2013; Devi et al., 2014; Longvah et al., 2017). This enables it to deliver multiple benefits in terms of health compared with major cereals (Saleh et al., 2013). Finger millet has been identified as one of the “future smart food crops” by FAO (Li and Siddique, 2018) because of its nutrient-dense and climate-resilient features; moreover, it can produce a reasonable yield at a relatively low cost of cultivation (Gupta et al., 2017). Finger millet grains are highly resistant to pest attacks and can be stored for long (Iyengar et al., 1945; Mgonja et al., 2007) and provide nutritional support to countries in the developing world (Mgonja et al., 2007; Gupta et al., 2017).

Germplasm is the basic raw material to drive any crop improvement. Its genetic characterization can lead to exploring the variation in germplasm. The great diversity in finger millet comes from its gene pool, including different races and subraces of cultivated species and nine genus *Eleusine* species (Sood et al., 2016; Mirza and Marla, 2019; Vetriventhan et al., 2020). While a large number of finger millet germplasms have been collected and conserved in repositories worldwide, a small fraction has been exploited for economically important traits. The genebank at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, conserves over 7,500 finger millet germplasms (<http://genebank.icrisat.org>). To enhance the use of diversity in crop improvement, core and mini-core collections were established in finger millet (Upadhyaya et al., 2006, 2010). For this study, we formed a large germplasm diversity representative subset of 314 accessions, originating from 23 countries. The subset included all the mini-core accessions, trait-specific sources identified from the core collection, selections by breeders, and recently assembled germplasm and elite breeding lines. Assessing the diversity and

structure of this subset is important for its effective utilization in genomic-assisted crop improvement.

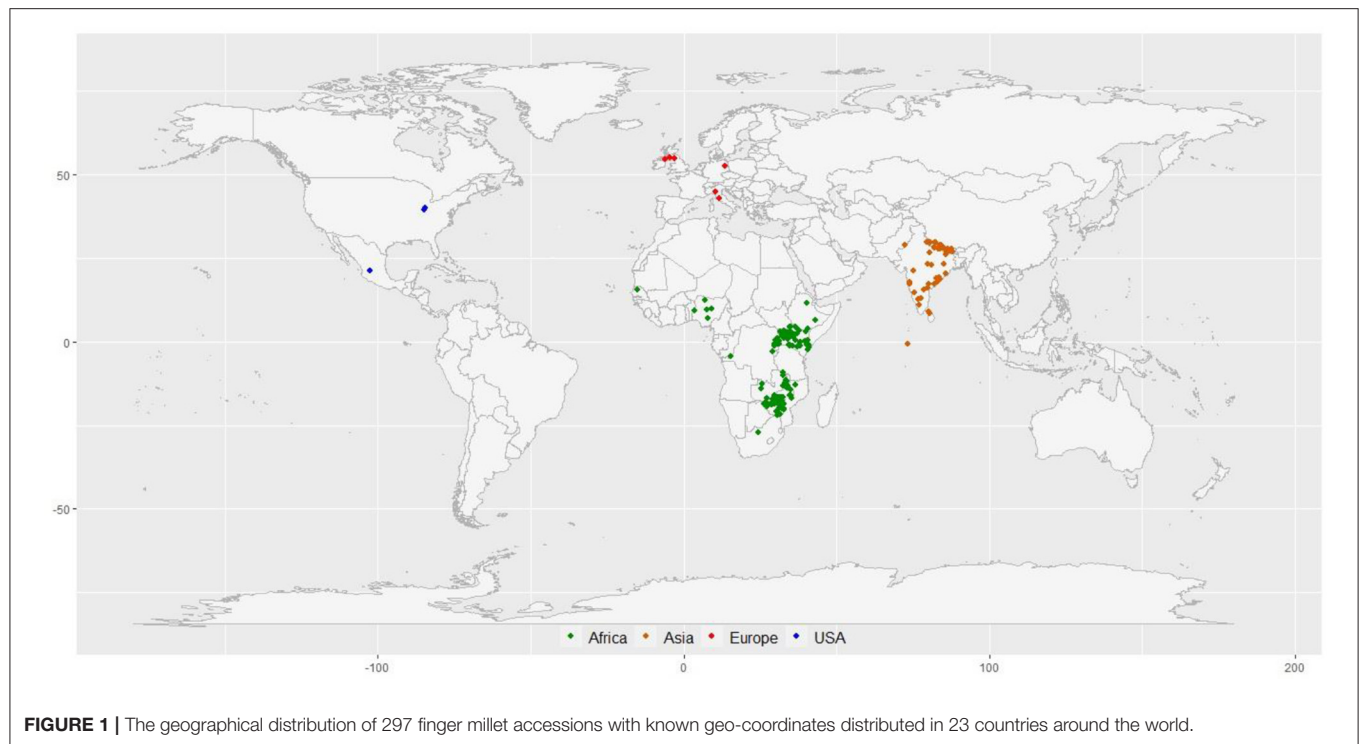
Although nutritional potential and climate resilient features of finger millet are being documented at the global level, it continues to be an orphan crop, lacking genomic information that can be used in crop improvement (Sood et al., 2016). The recent publication of the genome of finger millet (Hittalmani et al., 2017; Hatakeyama et al., 2018) has opened up opportunities to expand genome-level knowledge. Characterization of finger millet germplasm has been assessed, using morphological, and several molecular markers, such as RAPD (Fakrudin et al., 2004; Babu et al., 2007; Das and Misra, 2010; Gupta et al., 2010; Kumari and Pande, 2010; Panwar et al., 2010b; Singh and Kumar, 2010; Ramakrishnan et al., 2016b), ISSR (Salimath et al., 1995; Gupta et al., 2010), SSR (Srinivasachary et al., 2007; Dida et al., 2008; Panwar et al., 2010a,b; Bharathi, 2011; Kumar et al., 2012; Arya et al., 2013; Babu et al., 2014a, 2017; Kalyana Babu et al., 2014; Gimode et al., 2016; Pandian et al., 2018), and EST-SSR (Arya et al., 2009; Naga et al., 2012; Babu et al., 2014b; Bwalya et al., 2020). A few SNP-based studies too were attempted (Gimode et al., 2016; Kumar et al., 2016) to conduct GWAS studies for major agronomic and nutritional traits (Sharma et al., 2018; Puranik et al., 2020; Tiwari et al., 2020). Recent developments in next-generation sequencing have enabled the rapid genotyping of a larger number of germplasms, with a high number of marker loci at low cost. Recently, DArTseq-based SNPs have been used in many crops, especially in orphan crops and tree species where only limited genomic resources are available (Evanno et al., 2005; Kilian et al., 2012; Edet et al., 2018; O'Connor et al., 2019). DArTseq combines DArT with genotyping-by-sequencing technology, with the advantages of better genome coverage, high reproducible markers, and low cost of high throughput genotyping (Allan et al., 2020). This study aims to (i) assess the genetic diversity and population structure of a finger millet global diversity panel, (ii) assess the relationship among races, regions, landraces, and breeding lines, and (iii) identify the most diverse accessions. The information generated from this study can support understanding of the population structure and genetic diversity of finger millet germplasm, utilize the novel diversity to broaden the genetic base, and also accelerate genomics-assisted finger millet improvement.

MATERIALS AND METHODS

Plant Materials

A diversity panel of 314 finger millet accessions originating from 23 different countries was constituted for this study (Figure 1; Supplementary Table 1). These accessions represent four geographical regions in the world: Africa (160), Asia (136), Europe (6), North America (3), and unidentified origin (9). The panel represents all the races and subraces of finger millet: *vulgaris* (202 accessions), *plana* (48 accessions), *elongata* (31 accessions), *compacta* (28 accessions), and some unclassified (5 accessions). Each race was further grouped into subraces, except *compacta* that have no subrace. The diversity panel consists of both landraces (264 accessions) and breeding material (50 accessions) (Supplementary Table 1).

Abbreviations: DArTseq, diversity array technology sequencing; SNP, single nucleotide polymorphism; MAF, minor allele frequency; PIC, polymorphism information content; GD, gene diversity; MRD, modified Roger's distance; PCoA, principal coordinates analysis; MCMC, Markov Chain Monte Carlo; *Fst*, fixation index; AMOVA, analysis of molecular variance; GWAS, genome-wide association studies; QTL, quantitative trait loci.



DNA Extraction and DArT Sequencing

A total of 314 samples (including 4 control varieties, GPU 26, MR 6, KMR 204, and VL 149) were used for genotyping. Eight seeds of each accession were randomly chosen to constitute a sample for DNA extraction and sequencing. The seed samples from the 2018 rainy season harvested material were cleaned properly to avoid contamination by dust particles. The seeds were then placed in each well of the PCR plate and tightly sealed to avoid contamination. DNA extraction and sequencing were done by DArT private limited (www.diversityarrays.com); details regarding DArT genotyping methods and procedures can be found at <http://www.diversityarrays.com/dart-application> and in Kilian et al. (2012). The finger millet germplasm used in this study was conserved in the genebank at ICRISAT, Hyderabad, and is available to researchers globally, following the standard material transfer agreement (SMTA).

DATA ANALYSIS

SNP Filtering

The SNP markers obtained from DArTseq were filtered with a maximum threshold of 95% reproducibility, <20% missing values for markers, and <50% missing values for each accession to obtain high-quality SNP markers, using dartR packages in R software (Gruber et al., 2018).

Genetic Diversity and Structure Analysis

Locus-based diversity estimates, such as minor allele frequency (MAF), polymorphism information content

(Botstein et al., 1980), expected heterozygosity [also known as “gene diversity” (GD)] (Nei, 1973), and Shannon information index, were calculated. The genetic distance matrix was calculated, following the modified Roger’s distance (MRD) method (Wright, 1978), and a dendrogram was constructed, using the ward. D2 agglomerative hierarchical clustering method (Murtagh and Legendre, 2014). Principal coordinates analysis (PCoA) was performed, using the distance matrix obtained from MRD. The analysis was performed by custom-scripted codes in R program v.3.6.0 using “Adegenet” (Jombart, 2008), “ade4” (Dray et al., 2007), and “cluster” (Maechler et al., 2019) packages. Population structure was assessed, using STRUCTURE v.2.3.4 software (Pritchard et al., 2000), using an admixture model, with a K-value ranging from 1 to 8, with three independent runs. Burn-in time and Markov Chain Monte Carlo (MCMC) were set up to 10,000/50,000 iterations for each run. The optimal number of genetic groups (K) was determined by STRUCTURE HARVESTER (Earl and VonHoldt, 2012), based on the rate of change in the log probability of data between successive K-values (Evanno et al., 2005).

Population Differentiation and Genetic Diversity Indices

AMOVA was computed to ascertain the level of genetic differentiation within and among structure-defined populations, regions, races as well as a biological status, using 999 permutations, using R v.3.6.0 package “poppr” (Kamvar et al., 2014). Besides, indices that explain the diversity in the population were calculated. Genetic indices, such as Shannon’s Index (*I*) and gene diversity (*He*), were computed for regions, races, and

biological status using R v.3.6.0 package “*poppr*” (Kamvar et al., 2014).

RESULTS

SNP Filtering

A total of 46,336 DArTseq-based SNP markers were generated, and 33,884 polymorphic SNP markers were retained after filtering. Of the 314 finger millet accessions (310 accessions in the diversity panel and four controls), eight accessions (IE 2030, IE 2799, IE 2825, IE 3157, IE 3475, IE 3788, IE 6314, and VL 149) were removed for having >50% missing data; so, all the downstream analyses were carried out only on 306 accessions. The call rate ranged from 80 to 100%, of which 30,869 markers displayed a >85% call rate. Reproducibility ranged from 95 to 100%, and 88.80% of SNP markers showed >97% reproducibility.

Genetic Diversity Analysis

The polymorphism information content value ranged from 0.003 to 0.500 (maximum value for the biallelic marker) with a mean value of 0.110. The gene diversity for all loci in the entire set of germplasms ranged from 0.001 to 0.50, with an average of 0.114 (**Figure 2**). The frequency distribution of polymorphism information content (PIC), minor allele frequency (MAF), and gene diversity (H_e), considering landraces and breeding lines were depicted in the **Supplementary Figure 1**. It clearly showed the difference in the dynamics of gene diversity and minor allele frequency between landraces and breeding lines. The density graph showed low frequency of rare alleles in breeding lines as compared with landraces for the polymorphic SNPs. Similarly, the gene diversity distribution slope was narrower in breeding lines than in landraces. It shows that landraces hold greater diversity than breeding lines. Mean gene diversity was similar but slightly higher in the Asian (0.132) accessions compared with the African accessions (0.129). Among the races, *vulgaris* and *elongata* had higher gene diversity (0.144), followed by *compacta* (0.142), and the lowest gene diversity was in race *plana* (0.129). Likewise, landraces (0.144) showed greater gene diversity compared with breeding lines (0.117) (**Table 1**). The African (0.243) and Asian (0.242) accessions had similar Shannon Index values. While among races, *vulgaris* had a higher value (0.268), and *plana* had a low value (0.233). Similarly, landraces (0.273) had a higher Shannon Index than breeding lines (0.204) (**Table 1**).

Genetic distance was assessed by the modified Roger’s distance (MRD) method. The genetic distance among individuals ranged from 0.040 to 0.450 with a mean value of 0.301. Among regions, the highest mean genetic distance was found in the Asian accessions (0.286), followed by the African accessions (0.276). The genetic distance of accessions from Europe and North America was 0.247 and 0.217, respectively; however, the sample size in these two regions was low (<10) and was, therefore, ignored for further discussion. Among races, *elongata* had the highest genetic distance (0.326), followed by *vulgaris* (0.299) and *compacta* (0.280), and the lowest was in race *plana* (0.275). The average genetic diversity among landraces was 0.299, while breeding lines had a value of 0.267 (**Table 4**).

The genetic distance between the populations (region, race, and biological status) was measured by MRD and the pairwise estimates of *Fst*. Pairwise *Fst* estimates between Asia and Africa were significant (0.198), which indicates the presence of genetic differentiation between the geographical origin, and the genetic distance between the African and Asian was 0.327. Among races (**Table 2**), pairwise *Fst* estimate values were near zero, indicating no defined population among races. Between races, *vulgaris* and *elongata* had the greatest genetic distance (0.320), while *compacta* and *plana* had the lowest (0.281). The distance between *vulgaris* and *compacta* (0.294) and *vulgaris* and *plana* (0.302) was low compared with that between *elongata* and *compacta* (0.308) and *elongata* and *plana* (0.310). Significant genetic divergence and *Fst* were observed between landraces and breeding lines (0.313 and 0.136), respectively (**Table 2**).

Most Diverse Accessions

Considering the relatedness of the accessions, the most diverse individual accessions and pairs of accessions were identified (**Table 3**). Accession IE 6095, originating in Asia (Nepal), had the highest genetic distance with IE 2606 (0.450), originating in Africa (Malawi). IE 6095 also showed higher genetic distance with four accessions originating in Africa—IE 3399 (0.449), IE 8599 (0.449), IE 2869 (0.442), and IE 5291 (0.440). The 10 most diverse accessions were identified, of which seven (IE 5903, IE 6095, IE 6221, IE 6074, IE 6165, IE 5957, and IE 6059) were from Asia and three (IE 2869, IE 2645, and IE 2780) were from Africa. Accession IE 5903, originating in Asia, was the most distant from other accessions with an average of 0.369. Among African accessions, IE 2869 was the most divergent, with a mean of 0.368.

Population Structure

The ward.D2 agglomerative hierarchical clustering method, representing the relationship based on MRD, showed that 306 finger millet accessions were split into two distinct clusters (cluster I – Asia, and cluster II – Africa) based on geographical origin (**Table 4**; **Supplementary Table 2**). Cluster I consisted of 144 individuals, the majority from Asia (125), and cluster II comprised 162 individuals, of which 142 accessions were from Africa. Accessions originating in Europe and North America were present in both clusters. Accessions with an unidentified origin were mostly grouped in cluster II (seven accessions) than in cluster I (two accessions) (**Figure 3**). Principal coordinates analysis (PCoA) revealed the two distinct clusters that were based on geographical origin, suggesting that distinct genetic structure exist between African and Asian finger millet accessions. The results shown in **Figure 4** were in concordance with the clustering pattern of the dendrogram, and the first two principal coordinates account for 32.6% (22.3 and 10.3%) of the total observed variation.

The population structure among the 306 finger millet accessions was assessed with STRUCTURE v.2.3.4, and the results revealed the existence of two major populations (pop1 and pop2) according to a geographical region of origin ($K = 2$) (**Figures 5A,B**). The accessions in pop1 were mostly from Africa (87%) and those in pop2 mostly from Asia (88%). The accessions in pop1 were 142 from Africa, 8 from Asia, 3 from Europe,

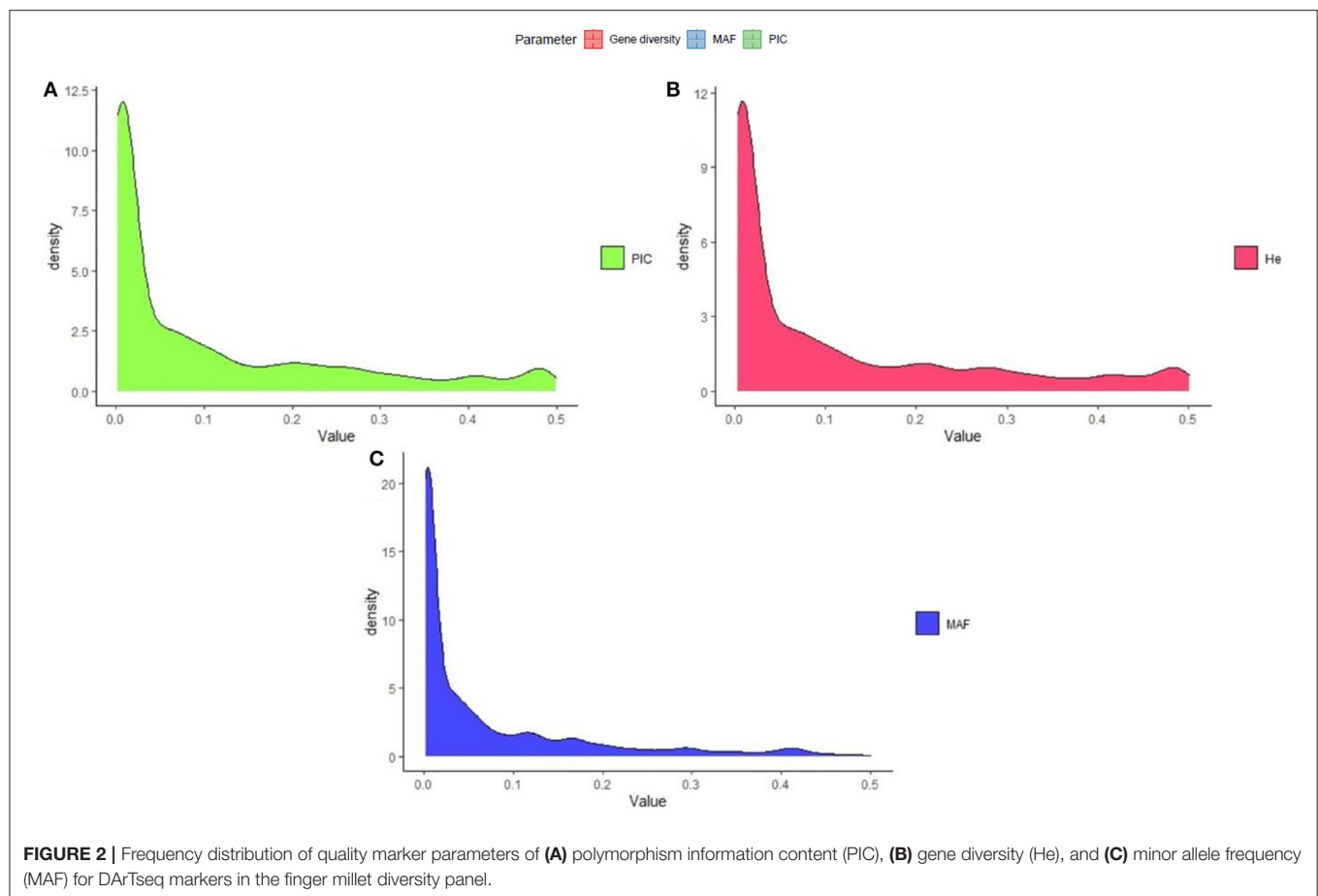


TABLE 1 | Shannon diversity index and gene diversity for the entire set, regions, races, and biological status of finger millet germplasm.

Population (number of accessions)	Shannon Index (<i>I</i>)	Gene diversity (<i>He</i>)
Overall mean	0.194 ± 0.01	0.114 ± 0.01
Region		
Africa (160)	0.243 ± 0.02	0.129 ± 0.01
Asia (136)	0.242 ± 0.03	0.132 ± 0.02
Races		
<i>Compacta</i> (28)	0.246 ± 0.06	0.142 ± 0.04
<i>Elongata</i> (31)	0.241 ± 0.05	0.144 ± 0.03
<i>Plana</i> (48)	0.233 ± 0.04	0.129 ± 0.03
<i>Vulgaris</i> (202)	0.268 ± 0.02	0.144 ± 0.01
Biological status		
Landraces (264)	0.273 ± 0.02	0.144 ± 0.01
Breeding lines (50)	0.204 ± 0.04	0.117 ± 0.03

TABLE 2 | Pairwise estimates of modified Roger's distance (MRD) (diagonal and upper diagonal) and *Fst* (lower diagonal) among the regions, races, and biological status of finger millet germplasm.

Population	MRD and <i>Fst</i>			
Region	Africa	Asia		
Africa	0.276	0.327		
Asia	0.198*	0.286		
Races	<i>Vulgaris</i>	<i>Compacta</i>	<i>Plana</i>	<i>Elongata</i>
<i>Vulgaris</i>	0.299	0.294	0.302	0.320
<i>Compacta</i>	0.024*	0.280	0.281	0.308
<i>Plana</i>	0.068*	0.022	0.275	0.310
<i>Elongata</i>	0.038*	0.027*	0.044*	0.326
Biological status	Landraces	Breeding lines		
Landraces	0.299	0.313		
Breeding status	0.136*	0.267		

*Significant at $p = 0.05$ level.

2 from North America, and 7 of unidentified origin. Pop2 accessions consisted of 13 from Africa, 123 from Asia, 2 from Europe, and 1 of unidentified origin. A total of five accessions (Asia-2; Europe-2; North America-1; and Unidentified origin-1) not grouped in any of the two populations were considered

admixture lines. Apart from that, the admixture of alleles between two subpopulations does exist, and pop2 had more admixture than pop1. The fixation index (*Fst*) estimated from STRUCTURE results for each of the two subpopulations was 0.548 and 0.622,

TABLE 3 | Genetically distant individuals and pairwise accessions identified using the modified Roger's distance method.

Accessions	Region	Country	Average Genetic distance from other accessions
IE 5903	Asia	Nepal	0.369
IE 2869	Africa	Zambia	0.368
IE 6095	Asia	Nepal	0.366
IE 2645	Africa	Malawi	0.360
IE 6221	Asia	Nepal	0.355
IE 6074	Asia	Nepal	0.355
IE 6165	Asia	Nepal	0.353
IE 5957	Asia	Nepal	0.353
IE 6059	Asia	Nepal	0.351
IE 2780	Africa	Malawi	0.351
Diverse pair of accessions			Genetic distance
IE 6095 Asia (Nepal) and IE 2606 Africa (Malawi)			0.450
IE 6095 Asia (Nepal) and IE 3399 Africa (Zimbabwe)			0.449
IE 6095 Asia (Nepal) and IE 8599 Africa (Kenya)			0.449
IE 5903 Asia (Nepal) and IE 8599 Africa (Kenya)			0.447
IE 5903 Asia (Nepal) and IE 3399 Africa (Zimbabwe)			0.446
IE 5903 Asia (Nepal) and IE 2606 Africa (Malawi)			0.444
IE 6165 Asia (Nepal) and IE 2606 Africa (Malawi)			0.442
IE 6095 Asia (Nepal) and IE 2869 Africa (Zambia)			0.442
IE 6221 Asia (Nepal) and IE 2606 Africa (Malawi)			0.441
IE 6221 Asia (Nepal) and IE 8599 Africa (Kenya)			0.441
IE 6165 Asia (Nepal) and IE 8599 Africa (Kenya)			0.441
IE 5903 Asia (Nepal) and IE 2645 Africa (Malawi)			0.440
IE 6095 Asia (Nepal) and IE 5291 Africa (Zimbabwe)			0.440
IE 6221 Asia (Nepal) and IE 3399 Africa (Zimbabwe)			0.439
IE 3130 Asia (India) and IE 2780 Africa (Malawi)			0.439

and the ancestry-inferred cluster proportion of the membership of the samples was 0.644 and 0.356. The gene diversity of the two subpopulations was 0.087 and 0.093 (**Table 5**).

In addition to population structure at $K = 2$, the STRUCTURE results showed two more peaks at $K = 6$ and $K = 8$ (**Figure 5A**). Interestingly, the populations at $K = 6$ and $K = 8$ were sub-structured from the major populations at $K = 2$ (**Supplementary Table 3**). In $K = 6$, accessions from Africa and Asia were divided into three groups each results in six populations ($K = 6$) (**Supplementary Figure 2**). In the case of $K = 8$, African and Asian accessions were divided into four and three groups, respectively, of which one population had all admixed individuals (six accessions) (**Supplementary Figure 3**).

Accessions were not clustered based on racial classification. For instance, 58% of race *vulgaris* was grouped into cluster I, while the remaining 42% was under cluster II (**Table 4**). Races such as *compacta* and *elongata* were grouped under both clusters (9 and 11 accessions in cluster I and 18 accessions each in cluster II, respectively). The majority of the accessions of race *plana* were grouped under cluster II (89%). Landraces were grouped in both clusters I and II (97 and 160 accessions, respectively). However, most of the breeding lines were grouped in cluster I (94%) since they were mainly developed from an Asian breeding program, particularly in India (**Table 4**).

Genetic Differentiation and Diversity Indices

Genetic divergence by analysis of molecular variance (AMOVA) (**Table 6**) for the STRUCTURE population revealed that 25.17% of the total variation was accounted for regional diversity and 37.90% of the variation within individuals. AMOVA on geographical regions revealed 42.36% variation between individuals within regions, 17.69% variation between regions, and 39.96% within individuals. When AMOVA analysis was assessed between races, most of the genetic variation (52.92%) was split into individuals within the race, while 4.60% reflected among races. AMOVA performed for biological status indicated that 13.24% of the total genetic variance was attributable to biological status diversity and 47.59% of the variation between individuals within a biological status.

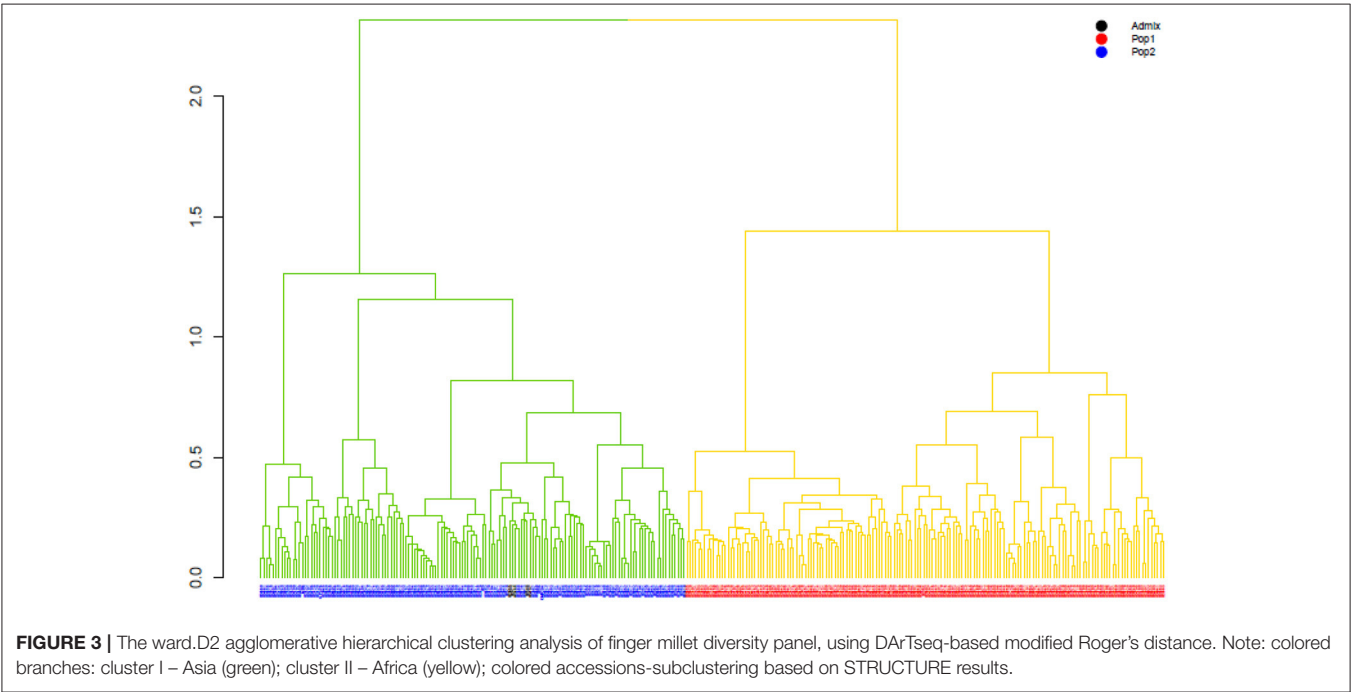
DISCUSSIONS

Finger Millet Diversity Panel

For this study, we established a diversity panel of 314 finger millet accessions, representing all the four races and their sub-races, originating from 23 countries. This set purposefully included the entire mini-core collection (80 accessions) (Upadhyaya et al., 2010) to capture the maximum diversity and all the

TABLE 4 | Dendrogram results and mean genetic distance of geographical regions, races, and biological status of the finger millet diversity panel.

		Cluster I (144 genotypes)	Cluster II (162 genotypes)	Average distance
Region	Africa	13	142	0.276
	Asia	125	8	0.286
	Europe	3	3	0.247
	North America	1	2	0.217
	Unidentified	2	7	0.271
Race	<i>Compacta</i>	9	18	0.280
	<i>Elongata</i>	11	18	0.326
	<i>Plana</i>	5	43	0.275
	<i>Vulgaris</i>	115	82	0.299
	Unclassified	4	1	0.291
Biological status	Landraces	97	160	0.299
	Breeding lines	47	2	0.267
Average distance	Within cluster	0.263	0.282	
	Between cluster		0.332	
	Overall range		0.040–0.450	
	Overall mean		0.301	



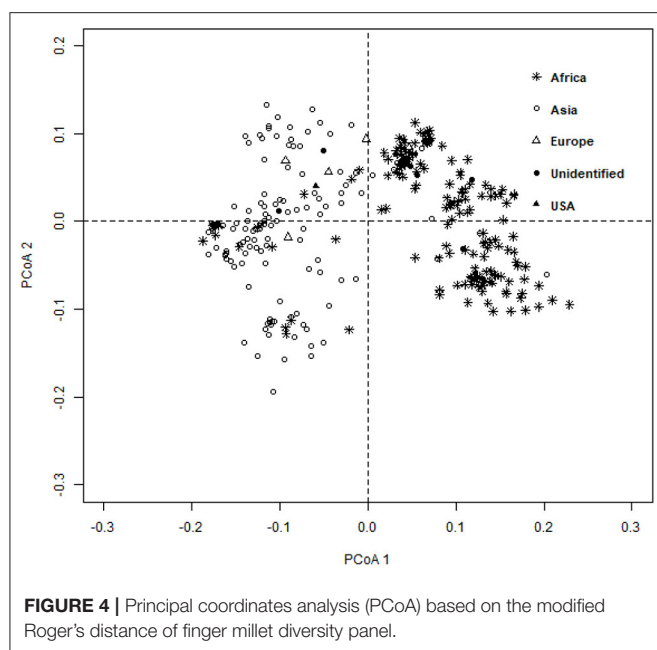
trait-specific sources identified in the core collection. Besides, new diversity from recently assembled accessions, selection by breeders, and advanced breeding lines or improved cultivars were included. This germplasm set can act as a potential diversity panel for phenotypic and genomic investigations for finger millet improvement.

Genetic Diversity and Allelic Richness

This study aimed at the genomic characterization of the finger millet diversity panel to understand diversity and population structure. Genotyping, using the DARtseq platform, provided

46,336 SNPs. Quality filters were applied to ensure high-quality markers to lessen the probability of false interpretation of downstream analysis (O'Connor et al., 2019), as genotyping errors occur irrespective of the DNA sequencing method used (Saunders et al., 2007). Locus-based diversity, population diversity, population structure, and genetic differentiation analyses were performed on the filtered high-quality DARtseq-based SNP markers (33,884).

The PIC of a marker is a good index to evaluate genetic diversity and estimate the level of genetic variation expressed by a particular marker. The average PIC value obtained from our



study was 0.110 (Figure 2), which was low compared with those obtained (0.150) for 59 cultivated accessions of finger millet, using SNP markers (Gimode et al., 2016), while the same study found a higher PIC value of 0.300 for wild finger millet. Many factors can influence PIC value, such as the breeding behavior of the species, size, and genetic diversity of the collection, the genotyping method, and the genomic location of markers (Singh et al., 2013; Chen et al., 2017). However, finger millet diversity studies have reported moderate to high informative PIC values (0.256 to 0.700), using SSR markers (Panwar et al., 2010a; Bharathi, 2011; Bheema Lingewara Reddy et al., 2011; Kumar et al., 2012; Babu et al., 2014a; Nirgude et al., 2014; Lee et al., 2017). In this study, the mean gene diversity and Shannon Index were 0.114 and 0.194, respectively, in an entire set (Table 1). Using SSR markers, gene diversity in finger millet has been reported to range from 0.35 to 0.57 (Dida et al., 2008; Babu et al., 2014b; Lee et al., 2017). Similarly, Upadhyaya et al. (2015) observed gene diversity of 0.28 in foxtail millet, using SNP markers; this was higher than finger millet in this study. Since very few diversity studies, using SNP markers, have been attempted in finger millet, more studies are needed for a better understanding of the finger millet germplasm diversity.

According to previous reports, African germplasm is more diverse compared with Asian germplasm (Dida et al., 2008; Panwar et al., 2010b; Bharathi, 2011; Arya et al., 2013; Kalyana Babu et al., 2014; Kumar et al., 2016; Ramakrishnan et al., 2016a; Babu et al., 2017). In this study, we found similar average gene diversity and genetic distance in African (0.129 and 0.276, respectively) and Asian (0.132 and 0.286, respectively) accessions, with slightly higher diversity in Asian than in African accessions. This could be due to the introduction and extensive utilization of African germplasm into the Asian finger millet breeding program, especially for introgression of blast resistance and for

new source of diversity, may result in equal or slightly higher diversity in Asian germplasm than in African germplasm. On *Fst* estimates, a value of more than 0.150 was considered as significant to differentiate between two populations (Frankham et al., 2010), as observed between Asia and Africa (Table 2). Moreover, only six accessions from Europe and three from North America were used in this study. Thus, the interpretation of results with fewer accessions from North America and Europe may lead to biased estimates in terms of genetic differentiation. Hence, the results of African and Asian accessions were taken into consideration for further discussion, and significant genetic differentiation was found between Asian and African accessions.

Cultivated finger millet is classified into four races (*elongata*, *compacta*, *vulgaris*, and *plana*) based on inflorescence morphology. Among the races, *elongata* had the highest average genetic distance (Table 4). The Shannon Index value indicated that race *vulgaris* was the commonly found ear type in finger millet with a mean genetic distance less than race *elongata*. However, the mean distance of races *compacta* and *plana* was less, leading to the conclusion that both contribute less variability to the germplasm compared with *elongata* and *vulgaris*. However, gene diversity demonstrated that race *vulgaris* and *elongata* hold more diversity than *plana* (Table 1). On *Fst* estimates, all races had values near zero, indicating no genetic differentiation among races but were morphologically distinct from one another in terms of panicle type as the races are primarily classified based on panicle morphology and shape (Hilu and DeWet, 1976) (Table 2). As expected, landraces had the highest gene diversity and genetic distance compared with breeding lines, as greater genetic variability in landraces and lower in breeding lines are indicative of a domestication bottleneck, and high selection pressure during the breeding process led to genetic erosion (Tables 1, 2). There was clear genetic differentiation between landraces and breeding lines (Table 2).

Population Structure

Geographical origin is a key determinant of population structure in finger millet germplasm (Dida et al., 2008; Kumar et al., 2016). Population structure, PCoA, and MRD clustering patterns of finger millet germplasms were largely consistent with previous classifications (Dida et al., 2008; Kumar et al., 2016; Puranik et al., 2020), which are based on geographical origin (Figures 3, 4, 5B). This demonstrated the genetic differentiation between finger millet accessions originating in Africa and Asia. A strong genetic structure existed in finger millet; thus the selection of accessions based on origin would be more diverse and effective in a finger millet crop improvement program. Cluster I from the dendrogram was in parallel with pop2 (Asia) from STRUCTURE results. Likewise, cluster II was in parallel with pop1 (Africa).

Although our results grouped individuals largely based on a geographical region, there were a few exceptions. For instance, a few Asian accessions were assembled into the African population and *vice versa*, in concurrence with previous reports (Dida et al., 2008; Ramakrishnan et al., 2016a; Sood et al., 2016). Eight Indian accessions (IE 3131, IE 4655, IE 4866, IE 5165, IE 5170, IE 5320, IE 954, and IE 5198) fell under pop1. Accession IE 3131 is Indaf 9,

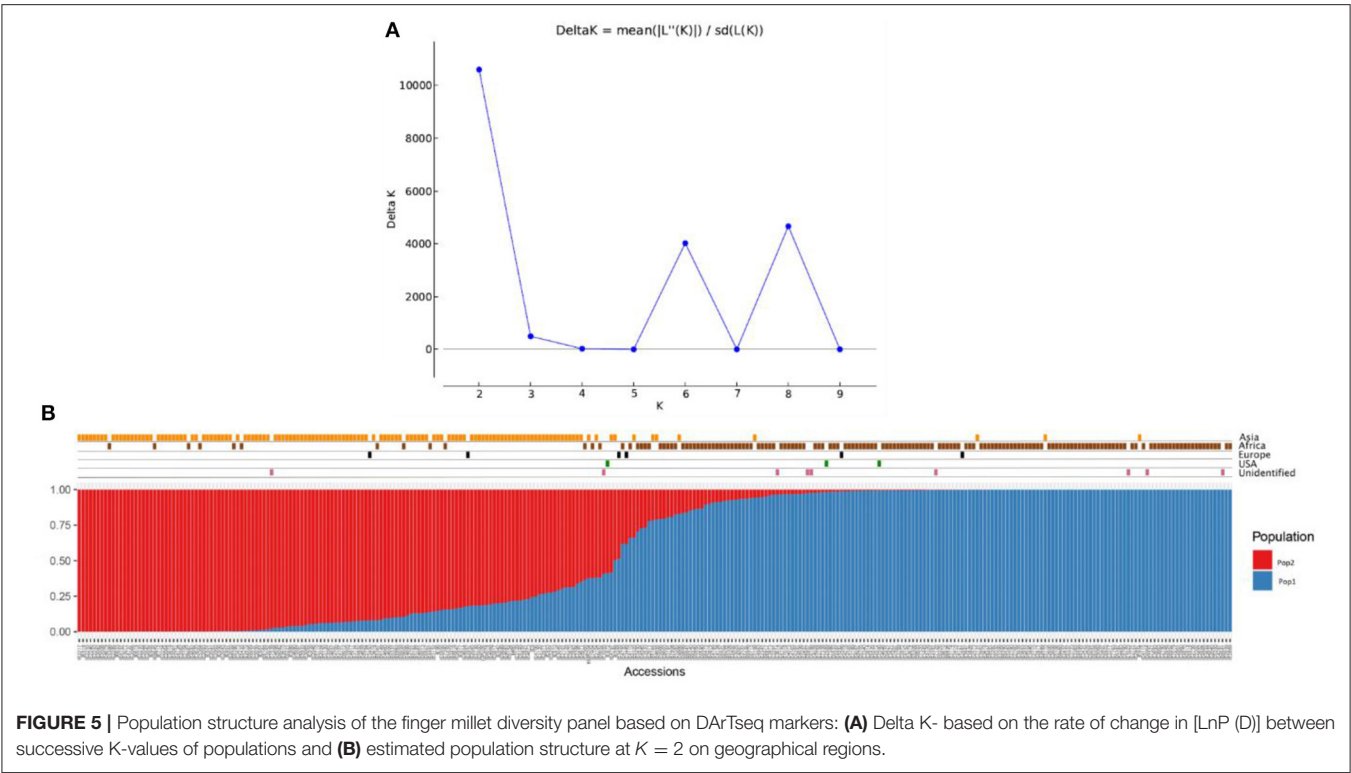


FIGURE 5 | Population structure analysis of the finger millet diversity panel based on DArTseq markers: **(A)** Delta K- based on the rate of change in [LnP (D)] between successive K-values of populations and **(B)** estimated population structure at K = 2 on geographical regions.

TABLE 5 | Gene diversity and *Fst* of the finger millet diversity panel for the two STRUCTURE -based populations.

K=2	Inferred clusters	<i>Fst</i>	Gene diversity	No. of accessions	Regional distribution of accessions
Pop1	0.644	0.548	0.087	162	Africa (142), Asia (8), Europe (3), North America (2), Unidentified (7)
Pop2	0.356	0.622	0.093	139	Africa (13), Asia (123), Europe (2), Unidentified (1)
	Admixture			5	Asia (2), Europe (1), North America (1), Unidentified (1)

and IE 4655 is GPU 13; both were released cultivars developed through hybridization between Indian and African germplasm; so, its placement under pop1 came as no surprise. IE 954 is a Co 4 variety developed through pure line selection of Palladam ragi landraces in Tamil Nadu, and the reason for its placement in the African population is not known. The remaining accessions were landraces with no information available on them. Likewise, 13 accessions of African origin were structured into the Asian population (pop2), of which five accessions (IE 4497, IE 6396, IE 6645, IE 6652, and IE 6667) were from Zimbabwe and eight accessions (IE 2658, IE 5364, IE 5367, IE 5388, IE 5433, IE 7390, IE 7404, and IE 8602) from Kenya. The possible hypothesis for this could be germplasm exchange between regions and their utilization in breeding programs.

Finger millet originated and was domesticated in about 5000 BC in western Uganda and the Ethiopian highlands and then spread to India in 3000 BC in the Western Ghats of India; thus, India is considered a secondary center of diversity for finger millet (Hilu and DeWet, 1976; Hilu et al., 1979; deWet et al., 1984). In this study, accessions from Europe and North America

were clustered in the Asian and African populations, for which there are two possibilities: either low sample size/insufficient samples from these regions to understand diversity, or that they were representing the native regions of the crop (Africa and Asia) and migrated to the respective regions in a breeding program, or through germplasm exchange/trading. Grouping accessions from different regions into a geographical origin show that finger millet germplasm mainly originated from Africa and were later introduced into Asia through a breeding process such as introduction and domestication, and exchange of germplasm led to the spread to other regions in the world. The five admixture accessions, IE 588 from India, IE 3455 from Europe, IE 872 from Mexico, IE4797 from the Maldives, and IE 1055 from an unidentified region, have been identified. The admixture is evidence of the continuous spill over of finger millet germplasm among different countries to date through breeding programs. There was a lower order of structure at K = 6 and K = 8, which signifies that there might be an existence of substructure within the geographical origin of finger millet accessions. However, K = 6 would be more

TABLE 6 | Analysis of molecular variance (AMOVA), using DArTseq markers among STRUCTURE populations, region, races, and biological status of the finger millet diversity panel.

Groups	Partitioning	Df	Mean Sq.	Sigma	Percentage of variation	P-value
STRUCTURE defined population	Between populations	1	211019.00	683.22	25.17	0.001
	Between individuals within a population	304	3034.08	1002.55	36.93	0.001
	Within individuals	306	1028.98	1028.98	37.90	0.001
Region	Between region	4	41762.67	455.52	17.69	0.001
	Between individuals within a region	301	3210.61	1090.80	42.36	0.001
	Within individuals	306	1029.00	1029.01	39.96	0.001
Race	Between races	3	12890.28	111.70	4.60	0.001
	Between individuals within a race	301	3594.29	1282.64	52.92	0.001
	Within individuals	306	1029.00	1029.01	42.46	0.001
Biological status	Between biological status	1	60752.64	347.62	13.24	0.001
	Between individuals within a biological status	304	3528.59	1249.79	47.59	0.001
	Within individuals	306	1029.00	1029.01	39.18	0.001

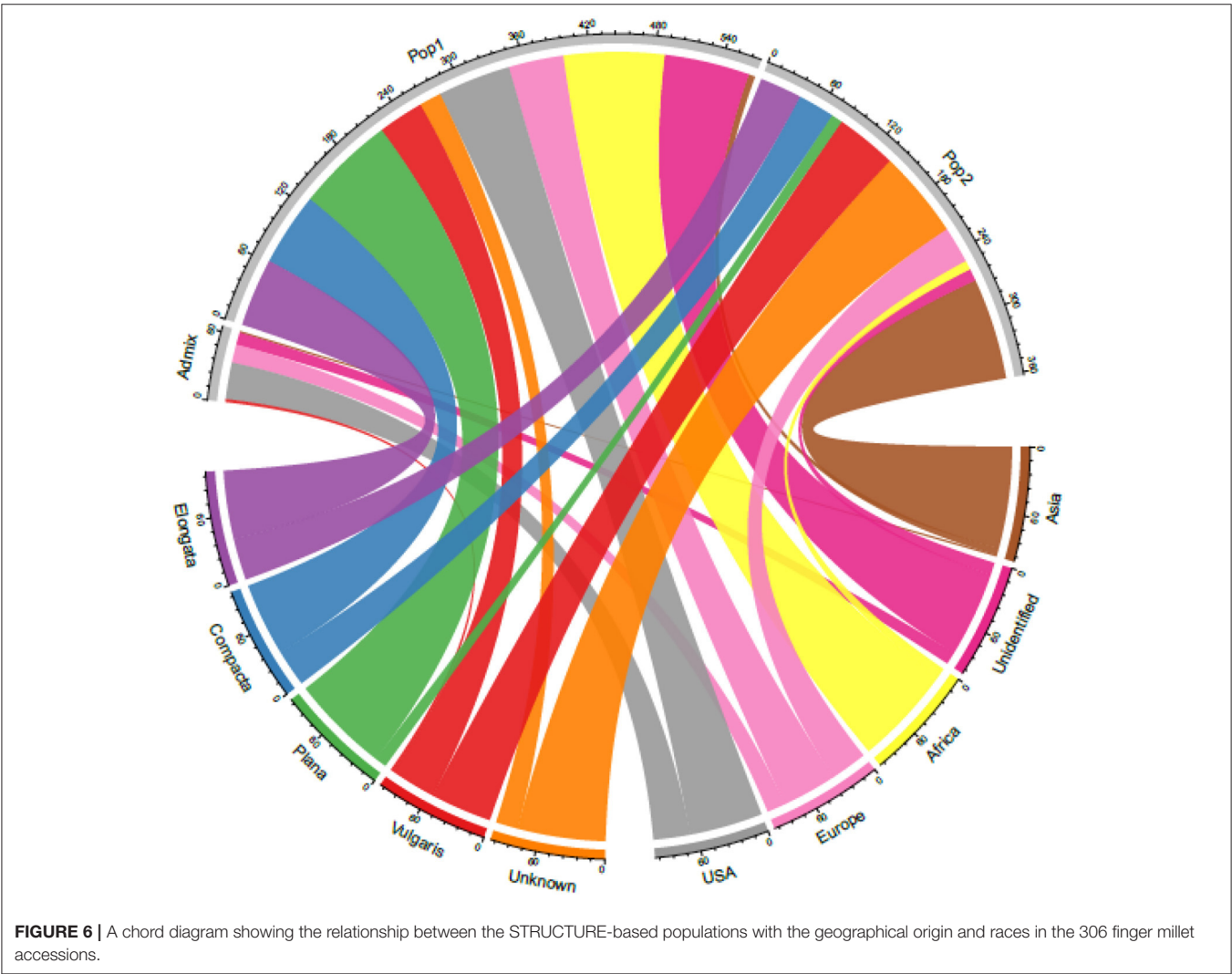


FIGURE 6 | A chord diagram showing the relationship between the STRUCTURE-based populations with the geographical origin and races in the 306 finger millet accessions.

appropriate to explain the population present in the finger millet accessions, while, in $K = 8$, one population had all admixed individuals (six accessions) (**Supplementary Figure 3**). Therefore, a structure at $K = 8$ might not be much informative compared with $K = 6$. In $K = 6$, accessions from Africa were structured into three populations (**Supplementary Figure 2**); of these, two populations consist of accessions from African lowland countries, such as Zimbabwe, Zambia, Malawi, and Tanzania. Accessions in the third population are mainly from African highland countries, such as Ethiopia, Kenya, and Uganda. Similarly, accessions from Asian origin were structured into three populations; of which, two were from India, while Nepalese accessions assembled into a separate population (**Supplementary Figure 2**). The population structure indicates the chronological order of domestication and introduction of finger millet in Africa and Asia. The African highland race is the most primitive form. Later, the lowland race was evolved from the highland race and subsequently introduced into India, where it formed into distinct gene pool over the period of time (Hilu and DeWet, 1976; deWet et al., 1984). Later, the African highland races were introduced into Nepal. The genetic differentiation and distinct structure between the region of origin (Africa and Asia) in finger millet are well known. In addition, this study gives new insights into the substructure within geographical origin of finger millet, which needs to be explored for crop improvement.

Although finger millet is classified into races and subraces, there are no studies on their impact on diversity and the relationship between races and geographical regions. Studies also have reported the lack of proper clustering among finger millet races (Bharathi, 2011; Kumar et al., 2016) (**Table 4**). In this study too, the accessions were not structured on the basis of races. This is unlike in foxtail millet, where accessions were structured mainly based on races and regions (Vetriventhan et al., 2014; Upadhyaya et al., 2015). In other small millets, such as proso millet (Vetriventhan et al., 2019), kodo millet (Johnson et al., 2019), and barnyard millet (Wallace et al., 2015), races are not a good indicator of genetic relatedness or to define population structure. The poor grouping among races could be because they were mainly classified based on panicle morphology and shape (deWet et al., 1984) and that the markers used in this study were not sufficient to capture the variation in genomic regions/gene spaces, encoding panicle morphology. On the contrary, the large number of high-quality SNPs does provide a possible capture of additional diversity that is not captured alone with panicle morphology-related gene complexes. This could also indicate that only panicle morphology-based racial differentiation is insufficient to capture the complete diversity in cultivated finger millet. A wide variation in the proportion of races in each cluster was observed, which is similar in the finger millet diversity study by Naik et al. (1993). The variation might be due to the predominance of accessions of race *vulgaris* (64%) compared with other races in the diversity panel and, also, in the entire germplasm of finger millet conserved at the ICRISAT Genebank. **Figure 6** explains the relationship between

the STRUCTURE population with the geographical regions and races.

Landraces dominate global finger millet germplasm collections conserved in different genebanks compared with improved or breeding lines, which comprise only ~3% of the total collection (Saha et al., 2016; Vetriventhan et al., 2020). The finger millet germplasm collection conserved at the ICRISAT genebank comprises 94% landraces, 3% breeding lines, and 3% wild relatives. The distribution of landraces was maximum in cluster II (62%) compared with cluster I (38%), where 94% of breeding lines were grouped into cluster I (**Table 4**). This demonstrates that Asian germplasm, in particular Indian accessions, has been subjected to more intense human selection in recent history after domestication. On the other hand, African germplasm remains unexplored and continues to be used mainly as a donor source in finger millet crop improvement programs, particularly in resistance breeding.

Genetic Differentiation and Diversity Indices

The AMOVA revealed the presence of genetic differentiation among STRUCTURE-derived populations, regions, and biological status (**Table 6**). The within-population AMOVA analysis explained most of the variance, indicating relatively unrestricted gene flow between regions, races, and biological status. Significant molecular variation was found between the STRUCTURE population and among regions, revealing the existence of genetic differentiation among the geographical regions. The races account for minimal variation in finger millet genetic diversity. For instance, landraces, and breeding lines have an impact on finger millet genetic variation. Altogether, high diversity was observed within populations and between individuals in a population, as reported in the finger millet previous study (Kalyana Babu et al., 2014; Pandian et al., 2018).

Diverse Accessions

A set of distantly related accessions was identified for use in hybridization and exploitation of hybrid vigor. Hybridization between distantly related individuals tends to yield superior hybrids; the introduction of new allelic combinations has been reported to result in increased heterozygosity (Fasahat et al., 2016). In India, the Indo-African hybridization program was set up to introduce new variability and resistance in finger millet varietal improvement (Dida et al., 2008). It resulted in the release of elite varieties in the country and gained popularity among Indian farmers. The released Indo-African varieties have been used as parents in successive hybridization programs. Consequently, diverse accessions identified from this study could be included in the finger millet varietal improvement program to introduce novel diversity to enable notable advancements in crop productivity. The diverse accessions from Africa include mostly those from African lowlands countries, such as Zimbabwe, Malawi, and Zambia, and Kenya from African highlands. Almost

all the top diverse Asian accessions identified were from Nepal. Nevertheless, India is a secondary center of diversity, and improvements within this gene pool will have limited opportunities (Dida et al., 2008). Therefore, the diverse accessions identified from African and Nepalese germplasm need to be explored for use in breeding programs to develop high-yielding cultivars.

CONCLUSIONS

This study provides a detailed understanding of the genetic differentiation between region, races, and biological status, using DArTseq-based SNP markers. The markers differentiated the population structure within the African and Asian regions. Genetic diversity was similar but slightly higher in the Asian accessions compared with the African accessions, probably due to greater integration of alleles from the African accessions through breeding. Finger millet races contribution to diversity was insignificant, and there was less association between geographical region and races; therefore, more attention should go towards the geographical region. The population structure identified in this study will aid in choosing appropriate statistical methods to perform genome-wide association studies (GWAS) and can be used to detect quantitative trait loci (QTLs)/gene in the finger millet population. Insights into finger millet diversity and population structure from this study will help breeders plan their breeding strategy to develop high-yielding cultivars with a broad genetic base for food, nutrition, and environmental security.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding authors.

REFERENCES

- Allan, V., Vetriventhan, M., Senthil, R., Geetha, S., Deshpande, S., Rathore, A., et al. (2020). Genome-wide dartseq genotyping and phenotypic based assessment of within and among accessions diversity and effective sample size in the diverse sorghum, pearl millet, and pigeonpea landraces. *Front. Plant Sci.* 11:1878. doi: 10.3389/fpls.2020.587426
- Arya, L., Verma, M., Gupta, V. K., and Karihaloo, J. L. (2009). Development of EST-SSRs in finger millet (*Eleusine coracana* ssp. *coracana*) and their transferability to pearl millet (*Pennisetum glaucum*). *J. Plant Biochem. Biotechnol.* 18, 97–100. doi: 10.1007/BF03263303
- Arya, L., Verma, M., Gupta, V. K., and Seetharam, A. (2013). Use of genomic and genic SSR markers for assessing genetic diversity and population structure in Indian and African finger millet (*Eleusine coracana* (L.) Gaertn.) germplasm. *Plant Syst. Evol.* 299, 1395–1401. doi: 10.1007/s00606-013-0822-x
- Babu, B. K., Agrawal, P. K., Pandey, D., and Kumar, A. (2014a). Comparative genomics and association mapping approaches for opaque2 modifier genes in finger millet accessions using genic, genomic and candidate gene-based simple sequence repeat markers. *Mol. Breed.* 34, 1261–1279. doi: 10.1007/s11032-014-0115-2
- Babu, B. K., Dinesh, P., Agrawal, P. K., Sood, S., Chandrashekara, C., Bhatt, J. C., et al. (2014b). Comparative genomics and association mapping approaches

AUTHOR CONTRIBUTIONS

MV and VCRA contributed to the conception and the design of the study. This work is part of Ph.D. thesis research of CBac. CBac supported student research as chairman. CBac, VA, and MV performed the statistical analysis. CBac and MV wrote the first draft of the manuscript. SD and RG provided the financial and technical support for this study. All the authors contributed to the revision of the manuscript, read, and approved the submitted version.

FUNDING

We acknowledge the funding support by the Department of Agriculture, Government of Karnataka for the project Genomics-Assisted Breeding for High Yielding and Climate-Resilient Finger Millet (ragi) Varieties/Hybrids and Promotion of Best Suitable Cultivars for Food and Nutritional Security in the Karnataka State of India, and the CGIAR genebank platform coordinated by Crop Trust for this study. This study was undertaken as part of the CGIAR Research Program on Grain Legumes and Dryland Cereals.

ACKNOWLEDGMENTS

The authors are grateful to Mr. Ramesh Reddy, Mr. Shanku Reddy, Mr. Raju Jadhav, and Mr. Manish Reddy for their help in field trials and also acknowledge Intertek India Private Limited for their help in genomics work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.692463/full#supplementary-material>

- for blast resistant genes in finger millet using SSRs. *PLoS ONE* 9:e99182. doi: 10.1371/journal.pone.0099182
- Babu, B. K., Senthil, N., Gomez, S. M., Biji, K. R., Rajendraprasad, N. S., Kumar, S. S., et al. (2007). Assessment of genetic diversity among finger millet (*Eleusine coracana* (L.) Gaertn.) accessions using molecular markers. *Genet. Resour. Crop Evol.* 54, 399–404. doi: 10.1007/s10722-006-0002-8
- Babu, B. K., Sood, S., Agrawal, P. K., Chandrashekara, C., Kumar, A., and Kumar, A. (2017). Molecular and phenotypic characterization of 149 finger millet accessions using microsatellite and agro-morphological markers. *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.* 87, 1217–1228. doi: 10.1007/s40011-015-0695-6
- Bharathi, A. (2011). *Phenotypic and genotypic diversity of global finger millet (Eleusine coracana (L.) Gaertn.) composite collection* (Ph.D.). Tamil Nadu Agricultural University, Coimbatore.
- Bheema Lingaswara Reddy, I. N., Srinivas Reddy, D., Lakshmi Narasu, M., and Sivaramakrishnan, S. (2011). Characterization of disease resistance gene homologues isolated from finger millet (*Eleusine coracana* L. Gaertn.). *Mol. Breed.* 27, 315–328. doi: 10.1007/s11032-010-9433-1
- Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.
- Bwalya, E. C., Marzougui, S., Mwangi, E., Wooseon, C., and Lee, M.-C. (2020). Genetic diversity analysis and population structure of some African and Asian Finger Millet (*Eleusine coracana* L.) accessions using expressed sequence tags –

- simple sequence repeat (EST-SSR) markers. *Int. J. Environ. Agric. Biotechnol.* 5, 643–646. doi: 10.22161/ijeab.53.17
- Chen, W., Hou, L., Zhang, Z., Pang, X., and Li, Y. (2017). Genetic diversity, population structure, and linkage disequilibrium of a core collection of ziziphous jujuba assessed with genome-wide SNPs developed by genotyping-by-sequencing and SSR markers. *Front. Plant Sci.* 8, 1–14. doi: 10.3389/fpls.2017.00575
- Chethan, S., and Malleshi, N. G. (2007). Finger millet polyphenols: optimization of extraction and the effect of pH on their stability. *Food Chem.* 105, 862–870. doi: 10.1016/j.foodchem.2007.02.012
- Das, S., and Misra, R. C. (2010). Assessment of genetic diversity among finger millet genotypes using RAPD markers. *Indian J. Agric. Res.* 44, 112–118.
- Devi, P. B., Vijayabharathi, R., Sathyabama, S., Malleshi, N. G., and Priyadarisini, V. B. (2014). Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. *J. Food Sci. Technol.* 51, 1021–1040. doi: 10.1007/s13197-011-0584-9
- deWet, J. M. J., Rao, K. E. P., Brink, D. E., and Mengesha, M. H. (1984). Systematics and evolution of *Eleusine coracana* (Gramineae). *Am. J. Bot.* 71:550. doi: 10.1002/j.1537-2197.1984.tb12540.x
- Dida, M. M., Wanyera, N., Harrison Dunn, M. L., Bennetzen, J. L., and Devos, K. M. (2008). Population structure and diversity in finger millet (*Eleusine coracana*) germplasm. *Trop. Plant Biol.* 1, 131–141. doi: 10.1007/s12042-008-9012-3
- Dray, S., Dufour, A.-B., and Chessel, D. (2007). The ade4 package - II: two-table and K-table methods. *R News* 7, 47–52.
- Earl, D. A., and VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7
- Edet, O. U., Gorafi, Y. S. A., Nasuda, S., and Tsujimoto, H. (2018). DARtSeq-based analysis of genomic relationships among species of tribe Triticeae. *Sci. Rep.* 8:16397. doi: 10.1038/s41598-018-34811-y
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Fakrudin, B., Kulkarni, R. S., Shashidhar, H. E., and Hittalmani, S. (2004). Genetic diversity assessment of finger millet, *Eleusine coracana* (Gaertn.), germplasm through RAPD analysis. *Plant Genet. Resour. Newsl.* 138, 50–54.
- Fasahat, P., Rajabi, A., Mohseni Rad, J., and Derera, J. (2016). Principles and utilization of combining ability in plant breeding. *Biometrics Biostat. Int. J.* 4, 1–22. doi: 10.15406/bbij.2016.04.00085
- Frankham, R., Ballou, J. D., and Briscoe, D. A. (2010). *Introduction to Conservation Genetics*. Cambridge, TN: Cambridge University Press.
- Gimode, D., Odeny, D. A., de Villiers, E. P., Wanyonyi, S., Dida, M. M., Mnene, E. E., et al. (2016). Identification of SNP and SSR Markers in finger millet using next generation sequencing technologies. *PLoS ONE* 11:e0159437. doi: 10.1371/journal.pone.0159437
- Gruber, B., Unmack, P. J., Berry, O. F., and Georges, A. (2018). dartr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* 18, 691–699. doi: 10.1111/1755-0998.12745
- Gupta, R., Verma, K., Joshi, D. C., Yadav, D., and Singh, M. (2010). Assessment of genetic relatedness among three varieties of finger millet with variable seed coat color using RAPD and ISSR markers. *Genet. Eng. Biotechnol. J.* 2, 1–9.
- Gupta, S. M., Arora, S., Mirza, N., Pande, A., Lata, C., Puranik, S., et al. (2017). Finger millet: A “Certain” Crop for an “Uncertain” future and a solution to food insecurity and hidden hunger under stressful environments. *Front. Plant Sci.* 8, 1–11. doi: 10.3389/fpls.2017.00643
- Hatakeyama, M., Aluri, S., Balachandran, M. T., Sivarajan, S. R., Patrignani, A., Grüter, S., et al. (2018). Multiple hybrid de novo genome assembly of finger millet, an orphan allotetraploid crop. *DNA Res.* 25, 39–47. doi: 10.1093/dnares/dsx036
- Hilu, K. W., de Wet, J. M. J., and Harlan, J. R. (1979). Archaeobotanical Studies of *Eleusine coracana* ssp. *coracana* (Finger Millet). *Am. J. Bot.* 66:330. doi: 10.1002/j.1537-2197.1979.tb06231.x
- Hilu, K. W., and DeWet, J. M. J. (1976). Racial Evolution in *Eleusine coracana* ssp. *coracana* (Finger millet). *Am. J. Bot.* 63, 1311–1318. doi: 10.1002/j.1537-2197.1976.tb13216.x
- Hittalmani, S., Mahesh, H. B., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y. R., et al. (2017). Genome and Transcriptome sequence of Finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18:465. doi: 10.1186/s12864-017-3850-z
- Iyengar, K. G., Doraisami, L. S., and Iyengar, R. S. (1945). Ragi (*Eleusine coracana*). *Mysore J. Agric. Sci.* 24, 33–33.
- Johnson, M., Deshpande, S., Vetriventhan, M., Upadhyaya, H. D., and Wallace, J. G. (2019). Genome-wide population structure analyses of three minor millets: kodo millet, little millet, and proso millet. *Plant Genome* 12:190021. doi: 10.3835/plantgenome2019.03.0021
- Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Kalyana Babu, B., Agrawal, P. K., Pandey, D., Jaiswal, J. P., and Kumar, A. (2014). Association mapping of agro-morphological characters among the global collection of finger millet genotypes using genomic SSR markers. *Mol. Biol. Rep.* 41, 5287–5297. doi: 10.1007/s11033-014-3400-6
- Kamvar, Z. N., Tabima, J. F., and Grunwald, N. J. (2014). Package “poppr” Type Package Title Genetic Analysis of Populations with Mixed Reproduction. *PeerJ* 1–14.
- Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., et al. (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. *Methods Mol. Biol.* 888, 67–89. doi: 10.1007/978-1-61779-870-2_5
- Kumar, A., Sharma, D., Tiwari, A., Jaiswal, J. P., Singh, N. K., and Sood, S. (2016). Genotyping-by-sequencing analysis for determining population structure of finger millet germplasm of diverse origins. *Plant Genome* 9, 1–15. doi: 10.3835/plantgenome2015.07.0058
- Kumar, A., Sharma, N., Panwar, P., and Gupta, A. K. (2012). Use of SSR, RAPD markers and protein profiles based analysis to differentiate *Eleusine coracana* genotypes differing in their protein content. *Mol. Biol. Rep.* 39, 4949–4960. doi: 10.1007/s11033-011-1291-3
- Kumari, K., and Pande, A. (2010). Study of genetic diversity in finger millet (*Eleusine coracana* L. Gaertn.) using RAPD markers. *African J. Biotechnol.* 9, 4542–4549.
- Lee, K. J., Yoon, M.-S., Shin, M.-J., Lee, J.-R., Cho, Y.-H., Lee, H.-S., et al. (2017). Development of SSR markers and their use in studying genetic diversity and population of finger millet (*Eleusine coracana* L. Gaertn.). *Plant Breed. Biotechnol.* 5, 183–191. doi: 10.9787/PBB.2017.5.3.183
- Li, X., and Siddique, K. H. M. (2018). Future smart food—rediscovering hidden treasures of neglected and underutilized species for zero hunger in Asia. *Bangkok* 4:36. doi: 10.18356/23b5f7ab-en
- Longvah, T., Ananthan, R., Bhaskarachary, K., and Venkaiah, K. (2017). *Indian Food Composition Tables*. National Institute of Nutrition (NIN). Hyderabad: Indian Council of Medical Research.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., Studer, M., et al. (2019). Package “cluster” for R: ‘Finding Groups in Data’: Cluster Analysis Extended. Available online at: <https://svn.r-project.org/R-packages/trunk/cluster>.
- Mgonja, M. A., Lenné, J. M., Manyasa, E., and Sreenivasaprasad, S. (2007). “Finger millet blast management in East Africa,” in *Creating Opportunities for Improving Production and Utilization of Finger Millet*, eds M. A. Mgonja, J. M. Lenné, E. Manyasa, and S. Sreenivasaprasad (Andhra Pradesh, India: ICRI), 196.
- Mirza, N., and Marla, S. S. (2019). “Finger millet (*Eleusine coracana* L. Gaertn.) breeding,” in *Advances in Plant Breeding Strategies: Cereals*, eds J. Al-Khayri, S. Jain, and D. Johnson (New York, NY: Springer, Cham), 83–132.
- Mundada, P. S., Nikam, T. D., Anil Kumar, S., Umdale, S. D., and Ahire, M. L. (2020). Morpho-physiological and biochemical responses of finger millet (*Eleusine coracana* (L.) Gaertn.) genotypes to PEG-induced osmotic stress. *Biocatal. Agric. Biotechnol.* 23:101488. doi: 10.1016/j.bcab.2019.101488
- Murtagh, F., and Legendre, P. (2014). Ward’s hierarchical agglomerative clustering method: which algorithms implement ward’s criterion? *J. Classif.* 31, 274–295. doi: 10.1007/s00357-014-9161-z
- Naga, B. L. R. I., Mangamoori, L. N., and Subramanyam, S. (2012). Identification and characterization of EST-SSRs in finger millet (*Eleusine coracana* (L.) Gaertn.). *J. Crop Sci. Biotechnol.* 15, 9–16. doi: 10.1007/s12892-011-0064-9
- Naik, B. J., Gowda, B. T. S., and Seetharam, A. (1993). “Pattern of variability in relation to domestication of finger millet in Africa and India,” in *Recent*

- Advances in Small Millets. Small Millets Workshop*, eds K. W. Riley, S. C. Gupta, A. Seetharam, and J. Moshanga (New York, NY: IBH Publishing Company), 347–364.
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci.* 70, 3321–3323. doi: 10.1073/pnas.70.12.3321
- Nirgude, M., Babu, B. K., Shambhavi, Y., Singh, U. M., Upadhyaya, H. D., and Kumar, A. (2014). Development and molecular characterization of genetic molecular markers for grain protein and calcium content in finger millet (*Eleusine coracana* (L.) Gaertn.). *Mol. Biol. Rep.* 41, 1189–1200. doi: 10.1007/s11033-013-2825-7
- O'Connor, K., Kilian, A., Hayes, B., Hardner, C., Nock, C., Baten, A., et al. (2019). Population structure, genetic diversity and linkage disequilibrium in a macadamia breeding population using SNP and silicoDART markers. *Tree Genet. Genomes* 15, 1–16. doi: 10.1007/s11295-019-1331-z
- Pandian, S., Satish, L., Rameshkumar, R., Muthuramalingam, P., Rency, A. S., Rathinapriya, P., et al. (2018). Analysis of population structure and genetic diversity in an exotic germplasm collection of *Eleusine coracana* (L.) Gaertn. using genic-SSR markers. *Gene* 653, 80–90. doi: 10.1016/j.gene.2018.02.018
- Panwar, P., Nath, M., Yadav, V. K., and Kumar, A. (2010a). Comparative evaluation of genetic diversity using RAPD, SSR and cytochrome P450 gene based markers with respect to calcium content in finger millet (*Eleusine coracana* L. Gaertn.). *J. Genet.* 89, 121–133. doi: 10.1007/s12041-010-0052-8
- Panwar, P., Saini, R. K., Sharma, N., Yadav, D., and Kumar, A. (2010b). Efficiency of RAPD, SSR and Cytochrome P450 gene based markers in accessing genetic variability amongst finger millet (*Eleusine coracana*) accessions. *Mol. Biol. Rep.* 37, 4075–4082. doi: 10.1007/s11033-010-0067-5
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945
- Puranik, S., Sahu, P. P., Beynon, S., Srivastava, R. K., Sehgal, D., et al. (2020). Genome-wide association mapping and comparative genomics identifies genomic regions governing grain nutritional traits in finger millet (*Eleusine coracana* L. Gaertn.) Societal Impact Statement. *Plants People Planet* 2, 649–662. doi: 10.1002/ppp3.10120
- Ramakrishnan, M., Antony Ceasar, S., Duraipandiyar, V., Al-Dhabi, N. A., and Ignacimuthu, S. (2016a). Assessment of genetic diversity, population structure and relationships in Indian and non-Indian genotypes of finger millet (*Eleusine coracana* (L.) Gaertn.) using genomic SSR markers. *Springerplus* 5:120. doi: 10.1186/s40064-015-1626-y
- Ramakrishnan, M., Ceasar, S. A., Duraipandiyar, V., Al-Dhabi, N. A., and Ignacimuthu, S. (2016b). Using molecular markers to assess the genetic diversity and population structure of finger millet (*Eleusine coracana* (L.) Gaertn.) from various geographical regions. *Genet. Resour. Crop Evol.* 63, 361–376. doi: 10.1007/s10722-015-0255-1
- Saha, D., Gowda, M. V. C., Arya, L., Verma, M., and Bansal, K. C. (2016). Genetic and genomic resources of small millets. *CRC. Crit. Rev. Plant Sci.* 35, 56–79. doi: 10.1080/07352689.2016.1147907
- Saleh, A. S. M., Zhang, Q., Chen, J., and Shen, Q. (2013). Millet grains: nutritional quality, processing, and potential health benefits. *Compr. Rev. Food Sci. Food Saf.* 12, 281–295. doi: 10.1111/1541-4337.12012
- Salimath, S. S., de Oliveira, A. C., Bennetzen, J. L., and Godwin, I. D. (1995). Assessment of genome origins and genetic diversity in the genus *Eleusine* with DNA markers. *Genome* 38, 757–763. doi: 10.1139/g95-096
- Saunders, I. W., Brohede, J., and Hannan, G. N. (2007). Estimating genotyping error rates from Mendelian errors in SNP array genotypes and their impact on inference. *Genomics* 90, 291–296. doi: 10.1016/j.ygeno.2007.05.011
- Sharma, D., Tiwari, A., Sood, S., Jamra, G., and Singh, N. K. (2018). Genome wide association mapping of agro- morphological traits among a diverse collection of finger millet (*Eleusine coracana* L.) genotypes using SNP markers. *PLoS One*, 9:e90972. doi: 10.1371/journal.pone.0199444
- Shobana, S., Krishnaswamy, K., Sudha, V., Malleshi, N. G., Anjana, R. M., Palaniappan, L., et al. (2013). “Finger Millet (Ragi, *Eleusine coracana* L.). a review of its nutritional properties, processing, and plausible health benefits,” in *Advances in Food and Nutrition Research* (Cambridge, MA: Academic Press Inc.), 1–39.
- Singh, N., Choudhury, D. R., Singh, A. K., Kumar, S., Srinivasan, K., Tyagi, R. K., et al. (2013). Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PLoS ONE* 8, 1–12. doi: 10.1371/journal.pone.0084136
- Singh, Y., and Kumar, J. (2010). Study of genomic fingerprints profile of Magnaporthe grisea from finger millet (*Eleusine Coracona*) by random amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR). *African J. Biotechnol.* 9, 7798–7804. doi: 10.5897/AJB10.963
- Sood, S., Kumar, A., Babu, B. K., Gaur, V. S., Pandey, D., Kant, L., et al. (2016). Gene discovery and advances in finger millet [*Eleusine coracana* (L.) Gaertn.] genomics—an important nutri-cereal of future. *Front. Plant Sci.* 7, 1634. doi: 10.3389/fpls.2016.01634
- Srinivasachary, Dida, M. M., Gale, M. D., and Devos, K. M. (2007). Comparative analyses reveal high levels of conserved colinearity between the finger millet and rice genomes. *Theor. Appl. Genet.* 115, 489–499. doi: 10.1007/s00122-007-0582-5
- Tiwari, A., Sharma, D., Sood, S., Jaiswal, J. P., Pachauri, S. P., Ramteke, P. W., et al. (2020). Genome-wide association mapping for seed protein content in finger millet (*Eleusine coracana*) global collection through genotyping by sequencing. *J. Cereal Sci.* 91:102888. doi: 10.1016/j.jcs.2019.102888
- Upadhyaya, H. D., Gowda, C. L. L., Pundir, R. P. S., Reddy, V. G., and Singh, S. (2006). Development of core subset of finger millet germplasm using geographical origin and data on 14 quantitative traits. *Genet. Resour. Crop Evol.* 53, 679–685. doi: 10.1007/s10722-004-3228-3
- Upadhyaya, H. D., Ramesh, S., Sharma, S., Singh, S. K., Varshney, S. K., Sarma, N. D. R. K., et al. (2011). Genetic diversity for grain nutrients contents in a core collection of finger millet (*Eleusine coracana* (L.) Gaertn.) germplasm. *F. Crop. Res.* 121, 42–52. doi: 10.1016/j.fcr.2010.11.017
- Upadhyaya, H. D., Sarma, N. D. R. K., Ravishankar, C. R., Albrecht, T., Narasimhudu, Y., Singh, S. K., et al. (2010). Developing a mini-core collection in finger millet using multilocation data. *Crop Sci.* 50, 1924–1931. doi: 10.2135/cropsci2009.11.0689
- Upadhyaya, H. D., Vetriventhan, M., Deshpande, S. P., Sivasubramani, S., Wallace, J. G., Buckler, E. S., et al. (2015). Population genetics and structure of a global foxtail millet germplasm collection. *Plant Genome* 8, 1–13. doi: 10.3835/plantgenome2015.07.0054
- Vetriventhan, M., Azevedo, V. C. R., Upadhyaya, H. D., and Naresh, D. (2019). Variability in the global Proso millet (*Panicum miliaceum* L.) Germplasm collection conserved at the ICRISAT Genebank. *Agric.* 9, 1–16. doi: 10.3390/agriculture9050112
- Vetriventhan, M., Azevedo, V. C. R., Upadhyaya, H. D., Nirmalakumari, A., Kane-Potaka, J., Anitha, S., et al. (2020). Genetic and genomic resources, and breeding for accelerating improvement of small millets: current status and future interventions. *Nucl.* 1:13. doi: 10.1007/s13237-020-00322-3
- Vetriventhan, M., Upadhyaya, H. D., Anandakumar, C. R., Senthilvel, S., Varshney, R. K., and Parzies, H. K. (2014). Population structure and linkage disequilibrium of ICRISAT foxtail millet (*Setaria italica* (L.) P. Beauv.) core collection. *Euphytica* 196, 423–435. doi: 10.1007/s10681-013-1044-6
- Wallace, J. G., Upadhyaya, H. D., Vetriventhan, M., Buckler, E. S., Hash, C. T., and Ramu, P. (2015). The genetic makeup of a global barnyard millet germplasm collection. *Plant Genome* 8, 1–7. doi: 10.3835/plantgenome2014.10.0067
- Wright, S. (1978). “Evolution and the genetics of populations,” in *Vol. 4 Variability Within and Among Natural Populations* (University of Chicago Press, Chicago. Scientific Research Publishing).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Backiyalakshmi, Vetriventhan, Deshpande, Babu, Allan, Naresh, Gupta and Azevedo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Applications of Multi-Omics Technologies for Crop Improvement

Yaodong Yang^{1*†}, Mumtaz Ali Saand^{1,2†}, Liyun Huang¹, Walid Badawy Abdelaal¹, Jun Zhang¹, Yi Wu¹, Jing Li¹, Muzafar Hussain Sirohi² and Fuyou Wang¹

¹ Hainan Key Laboratory of Tropical Oil Crops Biology/Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang, China, ² Department of Botany, Shah Abdul Latif University, Khairpur, Pakistan

OPEN ACCESS

Edited by:

Wolfram Weckwerth,
University of Vienna, Austria

Reviewed by:

Atsushi Fukushima,
RIKEN, Japan
Arindam Ghatak,
University of Vienna, Austria

*Correspondence:

Yaodong Yang
yyang@catas.cn

[†]These authors share first authorship

Specialty section:

This article was submitted to
Crop and Product Physiology,
a section of the journal
Frontiers in Plant Science

Received: 19 June 2020

Accepted: 06 August 2021

Published: 03 September 2021

Citation:

Yang Y, Saand MA, Huang L,
Abdelaal WB, Zhang J, Wu Y, Li J,
Sirohi MH and Wang F (2021)
Applications of Multi-Omics
Technologies for Crop Improvement.
Front. Plant Sci. 12:563953.
doi: 10.3389/fpls.2021.563953

Multiple “omics” approaches have emerged as successful technologies for plant systems over the last few decades. Advances in next-generation sequencing (NGS) have paved a way for a new generation of different omics, such as genomics, transcriptomics, and proteomics. However, metabolomics, ionomics, and phenomics have also been well-documented in crop science. Multi-omics approaches with high throughput techniques have played an important role in elucidating growth, senescence, yield, and the responses to biotic and abiotic stress in numerous crops. These omics approaches have been implemented in some important crops including wheat (*Triticum aestivum* L.), soybean (*Glycine max*), tomato (*Solanum lycopersicum*), barley (*Hordeum vulgare* L.), maize (*Zea mays* L.), millet (*Setaria italica* L.), cotton (*Gossypium hirsutum* L.), *Medicago truncatula*, and rice (*Oryza sativa* L.). The integration of functional genomics with other omics highlights the relationships between crop genomes and phenotypes under specific physiological and environmental conditions. The purpose of this review is to dissect the role and integration of multi-omics technologies for crop breeding science. We highlight the applications of various omics approaches, such as genomics, transcriptomics, proteomics, metabolomics, phenomics, and ionomics, and the implementation of robust methods to improve crop genetics and breeding science. Potential challenges that confront the integration of multi-omics with regard to the functional analysis of genes and their networks as well as the development of potential traits for crop improvement are discussed. The panomics platform allows for the integration of complex omics to construct models that can be used to predict complex traits. Systems biology integration with multi-omics datasets can enhance our understanding of molecular regulator networks for crop improvement. In this context, we suggest the integration of entire omics by employing the “phenotype to genotype” and “genotype to phenotype” concept. Hence, top-down (phenotype to genotype) and bottom-up (genotype to phenotype) model through integration of multi-omics with systems biology may be beneficial for crop breeding improvement under conditions of environmental stresses.

Keywords: multi-omics, crop sciences, genomics, transcriptomics, proteomics, ionomics, phenomics, panomics

INTRODUCTION

Various promising omics technologies have emerged over the last few decades. These omics-based approaches have proved themselves to be valuable for exploring the genetic and molecular basis of crop development through modifications in DNA, transcript levels, proteins, metabolites, and mineral nutrient against a backdrop of environmental and physiological stress responses (Muthamilarasan et al., 2019). Several omics approaches, such as genomics, mutagenomics, transcriptomics, proteomics, metabolomics, phenomics, and ionomics, have revealed each corresponding molecular biological facet integrated with plant systems (Salt et al., 2008; Houle et al., 2010; Talukdar and Sinjushin, 2015; Wu et al., 2017; Muthamilarasan et al., 2019). The advent of next-generation sequencing (NGS) technologies has led to high throughput and rapid data generation for genomes, epigenomes, transcriptomes, proteomes, metabolomes, and phenomes (Großkinsky et al., 2018). The integration of multiple omics approaches could elucidate gene functions and networks under conditions of physiological and environmental stress (Singh et al., 2013). Comprehensive multi-omics approaches with robust techniques have been used to identify and decipher essential components of stress responses, senescence, and yields in various economically important crops including wheat, soybean, and millet (Deshmukh et al., 2014; Talukdar and Sinjushin, 2015; Muthamilarasan and Prasad, 2017; Shah et al., 2018; Yadav et al., 2018).

In this review, we discuss multi-omics approaches, their applications, and anticipated implementations in crop science to improve crop yields and enhanced biotic and abiotic stress tolerance (**Figure 1**). We propose that the integration of entire omics approaches could provide a basis to improve genetic development, crop yields, crop breeding science, and crop resistance to physiological and environmental stress (**Figure 2**).

OMICS TECHNOLOGIES FOR CROP IMPROVEMENT

Genomics

Genomics deals with the study of genes and genomes and focuses on the structure, function, evolution, mapping, epigenomic, mutagenomic, and genome editing aspects (Muthamilarasan et al., 2019). Genomics can play an indispensable role in elucidating genetic variation, which may enhance crop breeding efficiency and subsequently result in the genetic improvement of crop species. Structural genomics encompasses sequence polymorphism and chromosomal organization and enables the construction of physical and genetic maps to identify traits of interest for plant biologists. In contrast, functional genomics provides insights into the functions of genes with regard to the regulation of the trait of interest. When epigenetic changes occur in the form of histone modifications, DNA, or small RNA methylations at the genomic level, the phenomenon is known as epigenomics. Mutagenomics deals with mutational events orchestrating genetic modification in mutant traits. However, pangenomics defined as sum of a core genome, shared by all individuals, plus a dispensable genome partially shared or

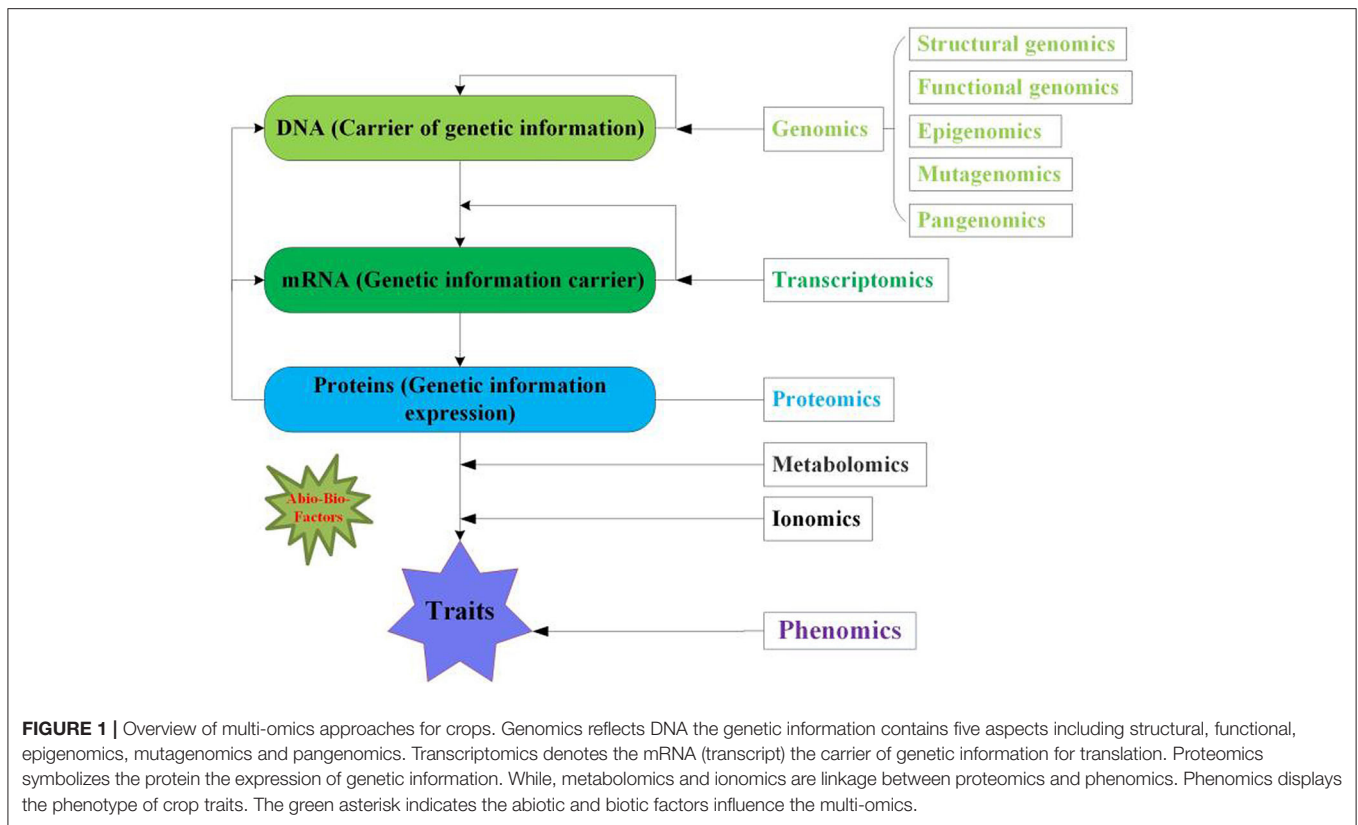
individual specific (Tettelin et al., 2005). Mutagenomics and pangenomics have emerged as recent omics approaches focused on mutagenesis and the pangenome in crop sciences, respectively (Golicz et al., 2016a; Goh, 2018; Muthamilarasan et al., 2019).

Structural Genomics

Structural genomics depends on molecular markers that may be useful for tagging and mapping genes of interest and their subsequent deployment in crop breeding programs. The marker techniques can be categorized into classes. The first one is non-PCR-based techniques which include restriction fragment length polymorphisms (RFLP). Restriction fragment length polymorphism detects DNA polymorphism through hybridizing labeled DNA probe to a Southern blot of DNA digested by restriction enzymes and resulting in differential DNA fragment profile (Agarwal et al., 2008). The second one is PCR-based techniques for markers such as, random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), and single nucleotide polymorphisms (SNPs) (Williams et al., 1990; Vos et al., 1995). The RAPD markers are PCR-based amplification of random DNA segments using single primer of arbitrary nucleotide sequence (Rabouam et al., 1999). Amplified fragment length polymorphisms is also the PCR-based technique which conducts selective PCR amplification of restriction fragments from a total digest of genomic DNA (Vos et al., 1995). Single nucleotide polymorphisms defined as single nucleotide variations in genome of an individual or an organism. The SNP may be performed through sequencing of genomic PCR products derived from varied individuals (Appleby et al., 2009). Whereas, the diversity arrays technology (DArT) a high-throughput technique which is based on microarray hybridization involving genotyping of numerous polymorphic loci spread over the genome (Jaccoud et al., 2001). The identification and usage of SNPs became possible with the advent of NGS.

Approaches utilized to understand and study the multiple traits in crops are quantitative trait loci (QTL) mapping and genome-wide association studies (GWAS). Quantitative trait loci mapping is a statistical method which assists in linking two types of data, i.e., complex phenotypes with genotypes. Molecular markers such as SNPs and AFLPs are commonly utilized for mapping QTLs, and then these may be correlated with observed phenotypic data (Kearsey, 1998; Challa and Neelapu, 2018). However, GWAS could identify variants associated with traits. Genome-wide association studies may also identify correlation between the genetic variants/phenotypes in a population of any organism based on SNPs in the sequence data (Challa and Neelapu, 2018).

The role of GWAS in genomics approaches is indispensable for enhancing the tolerance of crops to abiotic stress [e.g., the use of GWAS to evaluate how multiple abiotic stressors affect the oil content in sunflowers (*Helianthus annuus* L.) (Mangin et al., 2017)]. Previously, GWAS identified 48 QTLs related to the yield of maize crop under heat and water stress (Millet et al., 2016). Genome wide associations with environmental variables were used to predict the SNPs in sorghum (*Sorghum bicolor*) that were associated with drought stress (Lasky et al.,



2015). Another GWAS identified 213 unique genomic regions associated with drought tolerance in sorghum (Spindel et al., 2018). Genome-wide association studies have also identified the (drought resistance) DR-related loci in rice crop (Guo et al., 2018). Moreover, numerous SNPs associated with drought-responsive TFs have been identified using GWAS of maize crop (Shikha et al., 2017). In addition, structural variants (SVs) play a vital role in the genetic control of agronomically essential traits in crops. The association of SVs with agronomical traits has been reported in GWAS of *B. napus* (Gabur et al., 2018), maize (Lu et al., 2015), and soybean (Zhou et al., 2015).

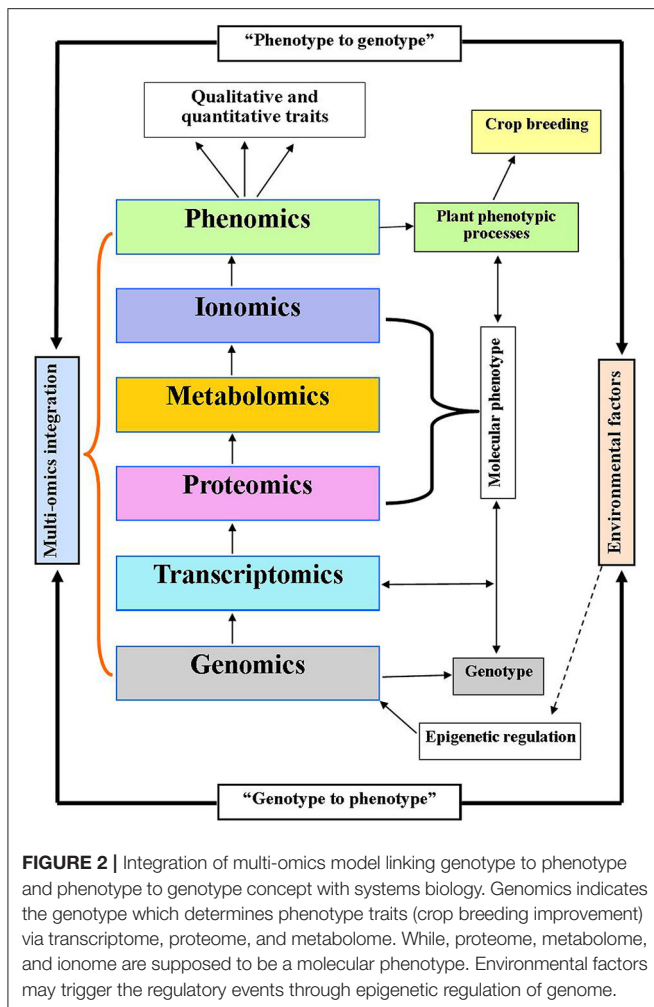
Breeders are now capable of enhancing hybrid breeding through marker-assisted selection (MAS) with genotyping-by-sequencing (GBS) to increase crop quality and yield (He et al., 2014). Multiparent mapping, in particular multiparent advanced generation intercrosses (MAGIC) and nested association mapping (NAM) in model plants and crops (Yu et al., 2008; Kover et al., 2009), has been able to expose the large amount of phenotypic diversity that may be achieved through experimental studies. The MAGIC population is ideal for breeding improvement. Analyses of the relationships between genotypes and phenotypes are able to identify QTLs that may be subsequently authenticated utilizing functional genomics approaches.

Functional Genomics and Mutagenomics

The vast resources and information provided through structural genomics will ultimately be utilized by functional

genomics. Functional genomics refers to development of global experimental approaches to assess the function of gene (Hieter and Boguski, 1997). Numerous biotechnological tools have been developed to identify and isolate genes of interest, to clone and characterize those genes, and to overexpress or knock-out lines for functional transgenic analyses (Muthamilarasan et al., 2019). Prior to genome sequencing approaches, the identification of candidate genes involved arduous procedures including suppression subtractive hybridization (SSH), expressed sequence tag (EST), and cDNA-AFLP-sequencing. Subsequently, the tediousness of these approaches decreased with the introduction of NGS (Muthamilarasan et al., 2019).

The access to crop genome sequencing has identified genes that play their role in disease resistance, stress resistance, and yield determination. Furthermore, authentic genome engineering has been envisaged to improve crops by utilizing genome editing tools such as the clustered regularly interspaced short palindromic repeats (CRISPR/Cas9 system) and transcription activator-like effector nuclease (TALEN) (Rinaldo and Ayliffe, 2015). Genome editing tools without the insertions of foreign DNA could possibly enhance yield via the introduction of pest and disease resistance in genetically modified crops. Using TALEN and CRISPR/Cas9 technologies, a bread wheat mildew resistance locus o (TaMlo) mutant was generated (Wang et al., 2014). Similarly, the same technique was adopted with tomato crop to create an SIMlo mutant (Nekrasov et al., 2017). Genome editing with the CRISPR/Cas9 system has already been reported for numerous important crops



including soybean, rice, maize, and sorghum (Jiang et al., 2013; Lawrenson et al., 2015; Li et al., 2015; Svtashev et al., 2015). Virus induced gene silencing (VIGS) is a reverse genetic technique to analyze the functions of genes that manifest in tomato crop in response to biotic and abiotic stress (Saand et al., 2015). Through comparative genomics, various mutants have been identified that are related to crop growth, development, and stress tolerance in rice, maize, wheat, and barley (Talukdar and Sinjushin, 2015). A soybean mutation has been found to alter the transcriptomic profiling of GmNARK (*Glycine max* leucine-rich repeat receptor kinase) rhizobia independent nodulation through the jasmonate pathway (Pathan and Sleper, 2008).

Mutagenomics emerged as a modern omics approach which enables to study mutational events orchestrating genetic modification in mutant traits. Such mutational events may be characterized by using high-throughput genomics technologies including serial analysis of gene expression (SAGE), high resolution melt (HRM), Targeted Induced Local Lesions IN Genomes (TILLING), and microarray analysis (Penna and Jain, 2017). Targeted Induced Local Lesions IN Genomes (McCallum et al., 2000) in functional genomics has characterized

mutagenesis and offers high throughput mutations in crops (Henikoff et al., 2004; Mba, 2013). Initially, TILLING technology was developed as a functional genomics strategy, but soon, it became a useful tool for crop breeding as an alternative to the transgenic approaches (Kurowska et al., 2011). The feasibility of using this technique has been documented in numerous essential crops, such as soybean wheat, rice, tomato, rapeseed (*Brassica napus*), and sunflower (Kurowska et al., 2011; Witzel et al., 2015). Thus, this technique has proved to be a potential method for functional genetics as well as a valuable tool to improve crop breeding (Chen et al., 2014). Mutants controlling the seed oil composition were screened through the reverse genetics technique TILLING (Knoll et al., 2011; Kumar et al., 2013). For example, mutants rich in oils, isoflavones, and oleic acids (FAD2 and FAD2-1B) have also been isolated in soybean crop (Pathan and Sleper, 2008). TILLING has also been applied to detect mutations in numerous crops including rice (Suzuki et al., 2008), maize (Till et al., 2004), wheat (Dong et al., 2009), barley (Caldwell et al., 2004), tomato (Minoia et al., 2010), and soybean (Cooper et al., 2008).

Several microarray analyses showed that plant mutagenesis could induce more transcriptomic changes compared with those due to transgene insertion (Varshney et al., 2010). Mutagenesis constitutes a vital technique to identify gene functions and develop countless agronomic traits with desirable variations (Henikoff et al., 2004; Varshney et al., 2010). Approximately 3,000 mutant varieties of various crops have been developed globally, of which 776 mutants ensure nutritional quality (Jain and Suprasanna, 2011). With improvements in functional, biological, and breeding tools, the mutagenomics has ensured induced mutagenesis in crops. However, various mutant traits have been identified in crop plants in perspective of global impact of mutation-derived varieties on food production and quality enhancement (Ahloowalia et al., 2004).

Mutagenomics through reverse genetic approaches have provided opportunity to silence and interrupt the candidate genes to investigate the function of gene. The specific reverse genetic techniques utilized to screen/induce mutations for crops that include, RNA Interference (RNAi) and (VIGS). When mutant alleles are not available, the reverse genetic techniques can be used to knockdown or silence the phenotype of gene, allowing analysis of gene function (Talukdar and Sinjushin, 2015). Furthermore, those reverse genetic approaches have been utilized to screen for mutations in wheat, rice, maize, barley, tomato, sunflower, cotton, chickpea (*Cicer arietinum* L.), pea (*Pisum sativum* L.), and soybean crops including RNAi and gene silencing technologies (Dwivedi et al., 2008; Gupta et al., 2008; Tomlekova, 2010).

As such, both functional genomics and mutagenomics have been found to be useful for improving crop growth, yield, and stress resistance.

Epigenomics

The epigenetics refers to heritable changes other than those in the DNA sequence. These epigenetic changes brought about through DNA methylation and post-translational modification (PTM) of histones (Strahl and Allis, 2000; Novik et al., 2002). The merger

of epigenetics and genomics is known as epigenomics which has arisen as new omics technique in order to understand the genetic regulation and its contribution to cellular growth and stress responses (Callinan and Feinberg, 2006). Unlike genomics, epigenomics may be influenced by environmental factors, including abiotic and biotic stress. Nevertheless, genome level studies could be conducted to analyze these epigenetic events at any developmental stage or to evaluate abnormalities due to plant disease (Muthamilarasan et al., 2019). The bisulfite sequencing technique can identify the DNA methylation status of the genome (Cokus et al., 2014) and has been successfully validated in tomato, maize, and soybean crops with regard to DNA methylation and chromatin regulated genes (Gent et al., 2013; González et al., 2013; Schmitz et al., 2013). The quantification of DNA methylation in the genome through a methylation-sensitive amplified polymorphism (MSAP) approach is common and has been performed in wheat and foxtail millet crops under salinity stress (Zhong et al., 2009; Pandey et al., 2017). Moreover, chromatin immunoprecipitation sequencing (ChIP-Seq) is a unique approach for the analysis of histone proteins and DNA methylation (van Dijk et al., 2010) and has been used in rice crop under drought stress (Zong et al., 2013). DNA methylation studies have been carried out with epigenome modifications associated with ripening in tomato and tissue cultured stable epigenome changes in rice crop (Stroud et al., 2013; Zhong et al., 2013). One epigenomic study found this approach to be beneficial for identifying histone modifications associated with photosynthesis in maize (Offermann et al., 2006).

Recently, an epigenome wide association study identified the MANTLED locus responsible for the mantled phenotype (hypomethylation) in the oil palm (*Elaeis guineensis*) (Ong-Abdullah et al., 2015). Karma (LINE) retrotransposon methylation was linked with normal fruit yield clones compared to mantled clones (Ong-Abdullah et al., 2015). Whole-genome bisulfite sequencing (WGBS) identified ncRNAs in cotton crop under drought stress (Lu et al., 2017). Taken together these data indicate that applications of epigenomics could play important role in crops improvement in response to environmental stresses.

Pangenomics

The pangenome concept refers to the full genomic makeup of a species, which can be divided into a set of core and dispensable genes. The sets of core genes are shared by all individuals, whereas, set of dispensable genes (also known as accessory genes) are individual-specific and/or present in some individuals but not all (Tettelin et al., 2005). Advancements in sequencing technology and analysis tools have made it possible to sequence several accessions of crop species (Golicz et al., 2016a). A wave of pangenomic studies in crops including rice (Schatz et al., 2014; Wang et al., 2018; Zhao et al., 2018), soybean (Li et al., 2014), wheat (Montenegro et al., 2017), maize (Hirsch et al., 2014), *Brassica napus* (Hurgobin et al., 2018), and *Brassica rapa* (Lin et al., 2014) have revealed that dispensable genes play important roles in maintaining crop diversity and improving quality. A pangenomic study using nine morphologically diverse *Brassica oleracea* varieties and a wild relative demonstrated that several variable genes were

annotated with functions related to major agronomic traits, such as glucosinolate metabolism, vitamin biosynthesis, and disease resistance (Golicz et al., 2016b). Numerous methodologies have been used to generate pangenomes in crops and their wild relatives, such as comparative *de novo*, iterative assembly and “map-to-pan.” Comparative *de novo* approaches have been conducted with soybean and rice and their wild relative in order to analyze the genetic variation and agronomic traits (Li et al., 2014; Zhao et al., 2018). While an iterative assembly approach was performed with *B. oleracea* (Golicz et al., 2016b), bread wheat (Montenegro et al., 2017), and *B. napus* (Hurgobin et al., 2018), and a “map-to-pan” approach was used with rice (Wang et al., 2018).

Pangenomic studies have recently been investigated to evaluate the genetic diversity of crop species. Comparatively dispensable genes tend to be more variable than core genes. For example, higher densities of SNPs and insertions/deletions (InDels) have been found in sets of dispensable genes when compared to those in sets of core genes in *Brachypodium distachyon* (Gordon et al., 2017), rice (Wang et al., 2018), and soybean (Li et al., 2014). The dispensable genes of a pangenome are determined by structural variation (Xu et al., 2012; Mace et al., 2013), and dispensable genomes have been found to be enriched with genes related to disease resistance in crops such as maize (Zuo et al., 2015) and rice (Fukuoka et al., 2009) and abiotic stress in barley (Francia et al., 2016) and sorghum (Magalhaes et al., 2007). Furthermore, pangenomics may be used to improve crops. Multiple crop wild relatives (CWRs) have been used in breeding programs specifically for their quantitative and adaptive traits. Traits associated with yield and its components (e.g., grain size) were subject to intensive selection during domestication and breeding improvement in crops including rice (Xiao et al., 1998), sorghum (Tao et al., 2017, 2018), and wheat (Huang et al., 2003). Several wild relatives have also been found to be able to contribute genes to improve traits, such as grain quality (Campbell et al., 2016) and biotic/abiotic stress resistance in crops (Zhang et al., 2006; Ram et al., 2007; Cao et al., 2011; Huang et al., 2013). Thus, pangenomic studies could be used to mine elite genes in CWRs for crop improvement.

Transcriptomics

Transcriptomics deals with transcriptome that refers to the complete set of RNA transcripts which are produced by genome of an organism in a cell or tissue (Raza et al., 2021). Transcriptome profiling is dynamic and has emerged as a promising technique to analyze gene expression in response to any stimuli over a certain time period (Duque et al., 2013; El-Metwally et al., 2014). This strategy helps the researcher to observe the differential expression of genes *in vitro* to understand the first layer function of a particular gene. Initially, transcriptome dynamics were analyzed using traditional profiling, cDNAs-AFLP, differential display-PCR (DD-PCR), and SSH, but these techniques provided low resolution (Nataraja et al., 2017). Soon after, the introduction of robust techniques made it possible for RNA expression profiling utilizing microarrays, digital gene expression profiling, NGS,

RNAseq, and SAGE (Kawahara et al., 2012; De Cremer et al., 2013; Duque et al., 2013).

Microarray analysis has revealed the differential expression of genes in soybean and barley during developmental and reproductive stages, respectively, under drought stress (Guo et al., 2009; Le et al., 2012). Similarly, the differential expression of genes was identified in soybean under dehydration stress using an Affymetrix GeneChip array (Khan et al., 2017). The expression of genes has been found to alter various TFs in *Arabidopsis*, soybean, and rice crops in response to abiotic stress (Xiong et al., 2002; Wohlbach et al., 2008). The novel TFs, Cys-2/His-2-type zinc finger (C2H2-ZF) TF and drought and salt tolerance (DST), were found to control stomatal aperture in response to salt and drought stress in rice crop (Huang et al., 2009). Another, study demonstrated the function of WRKY TFs in response to abiotic stress in wheat (Okay et al., 2014). Although phytohormones, non-coding RNAs, and small peptides regulate the expression of genes and are considered to be key components that execute gene functions in response to abiotic stress conditions in *Arabidopsis* and model crops including rice, tomato and wheat (Matsui et al., 2008; Chekanova, 2015; Bashir et al., 2019). Various phytohormone-independent abiotic stress responses are regulated by several TFs, such as DRE-/CRT-binding protein 2 (DREB2) and dehydration-responsive element/C-repeat (DRE/CRT), in rice crop (Todaka et al., 2015).

Transcriptome studies in sorghum revealed a set of differentially expressed (DE) genes in response to drought, heat, and osmotic stress as well as hormone treatment (Dugas et al., 2011; Johnson et al., 2014). Similarly, differential expression patterns of OsMADS genes were found in developing rice crop in response to drought stress (Jin et al., 2013). These transcriptome sequencing analyses have shown differential expression during growth and in response to stress and may be useful for functional analyses. Therefore, these reports demonstrate the role of transcriptomics in terms of stress responses and development for crops.

Novel advancements in transcriptomics have been brought about through *in situ* RNA-seq (i.e., *in situ* ligation), in which RNA is sequenced in living cells or tissues (Ke et al., 2013). Spatially resolved transcriptomics is another technique that detects gene expression with spatial information within cells or tissues to provide a comprehensive molecular description of physiological processes in organisms (Burgess, 2015). Numerous RNA-seq analyses have unveiled tissue-specific expression in response to abiotic and biotic stress in foxtail millet and sweet potato (*Ipomoea batatas* L.) crops (Qi et al., 2013; Hittalmani et al., 2014; Bonthala et al., 2016; Li et al., 2017). Total RNAseq has shown DE genes and SSR markers during the development of cowpea (*Vigna unguiculata* L. Walp.) crop (Chen et al., 2017). Thus, RNA-seq has proved to be one of the better techniques of transcriptomics to develop genic-SSR markers that can be linked to phenotypic traits connected with the candidate genes. Moreover, various studies utilizing RNA-seq technique have been conducted in rice, maize, and rapeseed oil to identify drought stress responsive genes (Kakumanu et al., 2012; Huang et al., 2014; Bhardwaj et al., 2015).

Comparative transcriptomics is another means to understand differential expression profiles in response to stress in different crop species. Comparative transcriptomic analysis has identified sixteen common genes in rice, wheat, and maize compared with those in switch grass in response to heat stress (Ding et al., 2013; Li et al., 2013). Comparative transcriptome and microarray analyses of biotic and abiotic stress and hormonal treatments have revealed multiple cross-talk pathways in cotton and potato (*Solanum tuberosum* L.) crops (Massa et al., 2013; Zhu et al., 2013). As such, these regulatory networks among stress tolerance genes might be beneficial for improving crops.

Recently, an alternative splicing (AS) transcriptomics approach was launched to generate multiple transcripts in response to abiotic stress conditions (Laloum et al., 2018). This method has been applied in crops including rice, maize, and sorghum in response to heat and drought stress (Zhang et al., 2015). Hence, AS transcriptomic analyses demonstrated the role for splicing factors controlling abiotic stress responses in crops. Collectively, these all transcriptomic techniques could play a vital role in the regulation of gene expression leading to the improvement of crop species.

Proteomics

Proteomics is a technique involved in the profiling of total expressed protein in an organism and is divided into four different parts including sequence, structural, functional, and expression proteomics (Mosa et al., 2017; Aizat and Hassan, 2018). Sequence proteomics determines the amino acid sequences that are usually identified sequentially utilizing high performance liquid chromatography (HPLC; Twyman, 2013). Structural proteomics deals with the structure of proteins to understand their putative functions. Structural proteomics can be analyzed through several approaches, such as computer based modeling, and experimental methods including nuclear magnetic resonance (NMR), crystallization, electron microscopy, and the X-ray diffraction of protein crystals (Sali et al., 2003; Woolfson, 2018). Functional proteomics determines the functions of a protein, and those functions are examined through various methods, such as yeast-one or two hybrids and protein microarray profiling (Lueong et al., 2014).

Advancements in protein extraction and separation have contributed to the rapid improvements of plant proteomic research, at both sample and genome-wide scales (Nakagami et al., 2012). Conventional proteomics are chromatography based techniques; which include exchange chromatography (IEC), size exclusion chromatography (SEC), and affinity chromatography. However, western blotting and enzyme-linked immunosorbent assay (ELISA) could be used for selective proteins analysis. Later, some more advanced techniques such as SDS-PAGE, two-dimensional gel electrophoresis (2-DE), and two-dimensional differential gel electrophoresis (2D-DIGE) were developed and used through gel based techniques for separation of proteins. Simultaneously, for rapid protein expression analysis the protein microarrays/chips have been devised for detection of small amount of protein sample. Moreover, stable isotope labeling with amino acids in cell culture (SILAC), Isotope-coded affinity tag (ICAT) labeling and isobaric tag for relative

and absolute quantitation (iTRAQ) have been developed as advanced techniques for quantitative proteomic analysis. Recently, two major high-throughput approaches including X-ray crystallography and NMR spectroscopy have been developed for three-dimensional structure determination of proteins that may be useful to understand the biological function of proteins (Aslam et al., 2017). Applications and importance from conventional to modern proteomic approaches have been discussed below.

Two-dimensional gel electrophoresis and SDS-PAGE are required to identify the proteins and measure the quantitative parameters of protein content, respectively (Eldakak et al., 2013). Henceforth, the identified proteins are used to analyze the molecular mass of peptides with mass-spectrometry (MS), ion trap-mass spectrometry (IT-MS), or liquid-chromatography (LC; Fournier et al., 2007). Additionally, the molecular weights of proteins have been identified using MALDI-TOF, electrospray ionization (ESI), and collision-induced dissociation (CID; Tanaka et al., 1988; McLuckey and Stephenson, 1998; Baggerman et al., 2005).

Functional proteomics approaches have identified ROS scavengers including quinone reductase, γ -glutamylcysteine synthetase, dehydrins, and dehydroascorbate reductase in tomato and sunflower crops (Shalata et al., 2001; Di-Baccio et al., 2004; Mittova et al., 2004). Meanwhile, molecular chaperones, such as heat shock proteins, have also been identified during proteome analyses in wheat and sugarcane (Demirevska et al., 2008; Jangpromma et al., 2010). Various drought sensitive and tolerant rice cultivars have been identified through comprehensive proteomics studies (Salekdeh et al., 2002; Rabello et al., 2008; Muthurajan et al., 2011; Maksup et al., 2014). Therefore, functional proteomics studies depict their significant role in crop defense response.

In quantitative proteomics, the iTRAQ method has demonstrated the differential expression of proteins in potato crop under abiotic stress (Liu et al., 2015). Recently, an iTRAQ-based comparative proteomic analysis of two coconut varieties identified numerous stress-responsive DEPs in two varieties of coconuts (Yang et al., 2020). In addition, iTRAQ-based proteomic analysis has provided new insights into somatic embryogenesis in cotton crop (Zhu et al., 2018). Thus, iTRAQ-based quantitative proteomic studies also play important role for crops against abiotic stresses.

A proteomics approach in response to the presence of pathogens was established with *Vitis* species (Basha et al., 2010). A post-iTRAQ-based comparative proteomic analysis was used to identify translational modifications (e.g., phosphorylation and ubiquitination) and protein-protein interactions that occur in biological or molecular mechanisms within cells. Phosphoproteomics intends to analyze protein phosphorylation through detecting phosphoproteins and their phosphorylated amino acid residues in a quantitative or qualitative manner (Mosa et al., 2017). In addition, phosphoproteomics has also been associated with protein functions, and thus it may play a role in the identification of pathways involved in various cell functions (Mosa et al., 2017). Phosphoproteomics, along with proteins have revealed large numbers of drought stress-related proteins in two

wheat crop cultivars (Zhang et al., 2014a). Further, proteomics and phosphoproteomics have been combined to investigate diverse functions in crops [e.g., wheat and grapevine (*Vitis vinifera* L.)] in response to phytoplasma and (*Septoria tritici*) fungal pathogen (Margaria et al., 2013; Yang et al., 2013). The wheat varieties against both drought and phytoplasma stresses showed resistant (Yang et al., 2013; Zhang et al., 2014a), whereas, grapevine was susceptible to phytoplasma infection (Margaria et al., 2013). Thus, phosphoproteomics could be helpful in order to identify resistant and/or susceptible crop cultivars against various stresses. Moreover, a combined proteomics and metabolomics approach with functional genomics in legumes has provided an understanding of the stress biology of these crops and the identification of molecular markers for legume smart breeding programs (Ramalingam et al., 2015). Hence, proteomics plays a vital role in deciphering functional mechanisms in crop science against diverse stresses and can help to improve crop yields.

Using LC-MS/MS, one proteomic study identified 75 differentially expressed proteins (DEPs) in a dehydration-sensitive chickpea cultivar (Subba et al., 2013). The majority of DEPs were involved in molecular chaperons, cell signaling, gene transcription, and regulation and ROS catabolic enzymes (Subba et al., 2013). Comparative proteomics and gene expression analyses using 2-DE along with LC-MS/MS have also identified DEPs associated with abiotic stress responses in chickpea (Arefian et al., 2019). By applying non-gel-based LC-MS/MS approaches, a large number of nodule proteins were identified in response to drought stress in *Medicago* spp. (Larrainzar et al., 2007). Later, the same method was used for the relative quantification of root nodule proteins in *Medicago* spp. (Larrainzar et al., 2009). Furthermore, using 2D-GE and ESI-LC-MS/MS approaches, numerous DEPs were identified in response to abiotic stress in legume crops, such as chickpea, common bean, and *M. truncatula* (Ramalingam et al., 2015). Several proteomics approaches, such as MALDI-TOF, SDS-PAGE, MS, 2-DE, and PMF have been applied in rapeseed, soybean, wheat, sugarcane, and cotton to determine stress-response pathways (Demirevska et al., 2008; Toorchi et al., 2009; Jangpromma et al., 2010; Nouri and Komatsu, 2010; Deeba et al., 2012; Mohammadi et al., 2012). These proteomics techniques have been implemented in plants under drought stress (Ghosh and Xu, 2014). However, 2-DE and SDS-PAGE proteomics techniques have been implemented in plants under drought stress (Ghosh and Xu, 2014). Importantly, numerous techniques (i.e., 2D-GE, MALDI-TOF, SDS-PAGE, ESI-IT- LC-MS/MS, and iTRAQ) used for cereal crops such as barley, maize, pearl millet, rice, sorghum, and wheat under drought stress response have been reviewed comprehensively by Ghatak et al. (2017). Hence, such proteomic studies revealed their role for crops in response to diverse abiotic stress conditions.

Metabolomics

Metabolomics defined as the comprehensive study of metabolites which participate in different cellular events in a biological system. However, metabolome denotes the complete set of metabolites synthesized via metabolic pathways in plant system

(Fiehn, 2002; Baharum and Azizan, 2018). Next-generation sequencing technologies have emerged as promising tools in order to understand the regulation of gene expression and molecular basis of cellular responses which occur in crops in response to biotic and abiotic stresses (Abdelrahman et al., 2018). However, metabolomics in combination with NGS has provided a basis to predict an initial metabolic network from a genome sequence of an organism (Weckwerth, 2011b). The genome sequencing approach (NGS) and quantification of metabolites (through MS) integrated the information in order to develop strategies for crop-improvement (Pandey et al., 2016).

Metabolites may be viewed as the end products of gene expression that display the biochemical phenotype of the cell (Weckwerth and Fiehn, 2002). Proteomics recognizes only gene products, whereas metabolomics may determine the expression of proteins metabolically and identify the biochemical processes that play important roles in gene functioning (Weckwerth and Morgenthal, 2005; Lindon and Nicholson, 2008).

Metabolites possess different chemical and physical properties; hence, separation and analytical techniques are required to generate metabolic profiles of a given plant sample (Jogaiah et al., 2013). Several analytical techniques have been implemented in plant systems to quantify metabolites including thin layer chromatography (TLC), gas/liquid-chromatography-mass spectrometry (GC/LC-MS), liquid chromatography-electrochemistry-mass spectrometry (LC-EC-MS), NMR, direct infusion mass spectrometry (DIMS), Fourier-transfer infrared (FT-IR), and capillary electrophoresis-liquid-chromatography mass spectrometry (CE-MS; Fiehn et al., 2000; Weckwerth, 2003; Moco et al., 2007; Allwood and Goodacre, 2010; Saito and Matsuda, 2010; Duque et al., 2013; Jogaiah et al., 2013). The CE-MS, GC-MS, LC-MS, and NMR techniques are the most frequently used in plant metabolomics (Fiehn, 2002; Kikuchi and Hirayama, 2007; Moco et al., 2007; Allwood and Goodacre, 2010; Weckwerth, 2010; Kim et al., 2011). These techniques depend on the selectivity, sensitivity, speed, and accuracy of the approach. NMR is fast and selective, although mass spectrometry techniques (CE-MS, GC-MS, and LC-MS) are suitably sensitive and selective but supposed to be time consuming (Sauter et al., 1991; Sumner et al., 2003).

Annotation and reporting of metabolomics data is an important in order to identify and analyze metabolites properly. However, recently, Alseekh et al. (2021) reported guidelines for annotation and quantification of LC/GC-MS-based metabolomics data reporting. Their recommended guidelines (i.e., sample preparation, sample replication and randomization, quantification, recovery and recombination, ion suppression, and peak misidentification) could be an effective tool/method for acquisition and reporting of metabolite data. Nonetheless, workflow for sampling, quenching, metabolite extraction, and storage has also been elucidated. The adopting certain recommendations may avoid misinterpretation of metabolite data and ensure the reporting transparency in LC/GC-MS-based metabolomics-derived data (Alseekh et al., 2021).

Plants have large chemically complex machinery that they employ as a major defense system against abiotic stress and pathogens. The mechanisms of plant metabolic responses to

stress depend on plant-stress or pathogen interactions. The pivotal role of metabolites in cereal crops, such as rice, maize, and barley, has been identified in the presence of various biotic stressors (Balmerl et al., 2013). The metabolic profiles of three rice varieties have identified several metabolites or biomarkers against the gall midge biotype 1 (GMB1) pathogen using GC-MS (Agarwal et al., 2014). Similarly, a number of metabolites were identified in rice crop in response to *Xanthomonas oryzae* pv. *oryzae* (Xoo; Sana et al., 2010). Another study demonstrated that the use of GC-MS could identify the accumulation of numerous metabolites in rice and barley crops against *Magnaporthe oryzae* (Parker et al., 2009). Meanwhile, phenylpropanoid and phenolic metabolites have been reported in wheat crop in response to biotic stress (Gunnaiah et al., 2012).

Metabolomics is particularly important in plant systems because plants produce more metabolites than either animals or microbes. The secondary metabolites produced by plants are helpful in responses to environmental stress. Thus, environmental metabolomics is a promising area in stress-physiology during that plant response to numerous abiotic stresses in relation to their metabolite changes (Brunetti et al., 2013; Viant and Sommer, 2013). Polyamine metabolites have been found in rice crop under drought stress conditions by applying GC-TOF-MS method (Do et al., 2013). In addition, a similar technique was used in rice, and the results identified salt tolerant cultivars (Liu et al., 2013; Gupta and De, 2017). Moreover, many metabolite analyses have been conducted in wheat, maize, tomato, and soybean crops in relation to drought, cold, and heat stress (Semel et al., 2007; Bowne et al., 2012; Silvente et al., 2012; Witt et al., 2012; Sun et al., 2016; Le et al., 2017; Paupiere et al., 2017). Several metabolomics techniques including LC/GC-MS, GC/EI-TOF-MS, HPLC, and NMR have been widely used in crop species such as, rice, tomato, maize and soybean in response to abiotic (drought, salt, oxidative, and temperature) and biotic stress conditions (Ghatak et al., 2018). Plant metabolite changes via certain pathways have been found to improve the nutritional value of genetically modified rice by the accumulation of β -carotene in the endosperm (Paine et al., 2005). Anthocyanin (a secondary metabolite) production was enhanced using metabolic engineering in tomato crops (Butelli et al., 2008). Therefore, metabolomics, coupled with other omics like genomics, transcriptomics, and proteomics, provide an integrated portrait of various functions ranging from the genome to metabolome as well as phenotypic characteristics (Weckwerth, 2011a). Strong correlations among these integrated omics have been identified in the responses of crops and plants to abiotic stress (Urano et al., 2010; Duque et al., 2013; Jogaiah et al., 2013).

Ionomics

Ionomics deals with the ionome, whereas ionome refers to the total mineral nutrient and trace elemental composition and represents the cellular inorganic components of plant systems (Salt et al., 2008; Satismruti et al., 2013). Ionomics comprises the quantitative measurement of the elemental composition of an organism and identifies the changes in mineral composition triggered by various physiological stimuli, genetic modifications, or developmental conditions. It is a dynamic

approach that is able to analyze the functions of genes and gene networks that characterize the ionome in response to physiological and environmental stress (Baxter, 2010). Ionomics acquisition by high throughput elemental profiling in plants has been conducted using different analytical tools including inductively coupled plasma-mass/optical emission spectroscopy (ICP-MS/OES), neutron activation analysis (NAA), and X-ray crystallography (Salt et al., 2008; Kumari et al., 2015). Inductively coupled plasma-mass spectroscopy technique is cheaper and can run hundreds of samples daily with excellent sensitivity to determine the elements. Nonetheless, ICP-OES may detect elements high throughput technique at the cost of some elements and sensitivity compared to ICP-MS. The XRF is faster to localize the elements but little cost than both ICP-MS/OES. Whereas, NAA is costly, slow and can detect more than 30 elements possible simultaneously (Salt et al., 2008).

In addition, the leaf ionome has been analyzed to identify the plant ionic regulatory networks involved in iron and phosphorus homeostasis (Baxter, 2015). Using ICP-MS, the leaf and grain ionomes have been analyzed to generate genetic maps, identify QTLs, and detect mineral element genetic diversification in rice crop (Norton et al., 2010; Zhang et al., 2014b; Pinson et al., 2015).

Moreover, seed ionome analysis has found differential gene expression, improved symbiotic responses to mycorrhizal fungi, and altered growth phenotypes under phosphate starvation in maize crop (Mascher et al., 2014). Single seed-based ionomic profiling in maize crop may be influenced by environmental and genetic factors that affect seed ionome accumulation (Baxter et al., 2014). Comprehensive, elemental profiling has revealed the QTLs responsible for grain mineral accumulation and yield in maize crop (Gu et al., 2015). Additionally, one ionomics study elucidated the relationships and responses of elements, minerals, and metabolites in barley under salt stress (Wu et al., 2013). Following this, the ionomic screening was performed in mutant lines of soybean crop which altered seed ionomic composition. They determined elemental concentration by applying ICP-MS method (Ziegler et al., 2013). Elemental profiling analysis has also been conducted in tomato cultivars to observe the concentrations of micro- and macro-nutrients under water stress (Sanchez-Rodriguez et al., 2010). Similarly, ionomic profiling has been performed to analyze the nutrient balance in some fruit species including kiwifruits, oranges, mangos, apples, and blueberry (Parent et al., 2013). Therefore, those ionomic studies suggest the important role for crop improvement and responses to various abiotic and biotic stimuli.

In light of these results, the integration of ionomics with other omics, such as genomics or metabolomics, could serve to identify potential genes and their networks that improve crop resistance in response to physiological and environmental stress (Singh et al., 2013; Wu et al., 2013; Huang and Salt, 2016; Guo et al., 2017).

Phenomics

Phenomics defined as the characterization of phenotypes through the acquisition of high-dimensional phenotypic data on an organism-wide scale (Houle et al., 2010). However, phenome

refers to the phenotype as a whole and plant phenome can be determined by genome (G), environment (E), and management (M) interactions (Gjuvsland et al., 2013; Großkinsky et al., 2018), thus phenomenon is also referred to as genotype–phenotype–enviotype (G–P–E) interactions (Zhao et al., 2019).

Precise phenotyping is very accurate for the gene and QTL mapping of particular traits of interest to identify their roles via forward and reverse phenomics applications for the genetic improvement in crop plants (Kumar et al., 2015b). In both cases, the “best of the best” or the “best varieties of the best” germplasm lines could be detected through automated high-throughput imaging technologies. These non-invasive imaging approaches allow for rapid phenotyping of the traits (phenes) through color imaging of the biomass, far infrared imaging of the canopy, lidar (light detection and ranging) to measure growth parameters, and magnetic resonance imaging to analyze root systems in crops (Finkel, 2009; Berger et al., 2010; Furbank and Tester, 2011). Furthermore, roots could be imaged in laboratories and greenhouses without damaging plant samples. For example electrical resistance tomography, electrical capacitance, X-ray computed tomography, and positron emission tomography were used to image root system for soil-grown plants and crops (McGrail et al., 2020). Red, green, blue (RGB) imaging, based on visible light, is a phenotyping tool used to estimate canopy and root systems (Großkinsky et al., 2018). Several studies have applied visible light imaging techniques through RGB setups in order to determine crop phenotyping parameters. For example, an RGB imaging setup scanned root system for *Triticum durum* grown in soil-filled rhizoboxes (Bodner et al., 2017), plant photosystem, and disease symptoms were assessed through RGB color-based imaging (Mahlein, 2016) and RGB digital imaging method was used to analyze plant shoots phenotyping under various stress responses (Humplík et al., 2015). Using infrared thermography, one study confirmed the role of stomatal conductance under salinity stress in young barley and wheat seedlings (Sirault et al., 2009). Chlorophyll fluorescence imaging has also been applied to screen for abiotic stress responses in tobacco, canola, and cotton crops using pulse amplitude modulated (PAM) instruments (Saranga et al., 2004; Baker, 2008). Furthermore, digital imaging methods have quantified boron toxicity under abiotic stress in form of the mapping of wheat and barley populations (Schnurbusch et al., 2010). With regard to the responses to biotic stress, similar method has been applied to detect and quantify the disease symptoms caused by pathogens in barley crop (Swarbrick et al., 2006; Chaerle et al., 2009). Thus, phenomics applications may play a vital role in order to evaluate the phenotypic parameters in crops under biotic and abiotic stress conditions.

Phenotyping techniques are more important for analyzing crops in the field compared to plants in either the laboratory or greenhouse. Multi and hyperspectral technologies may be utilized to determine various agronomic characteristics (Rascher and Pieruschka, 2008). Among these, the laser-induced fluorescence transient (LIFT) technique is one of the most robust for analyzing the photosynthetic efficiency of crops in the field (Pieruschka et al., 2010). Access to wireless sensor network aids in the phenotyping of crop traits and enables the

accumulation of valuable data for breeding science (Ruiz-Garcia et al., 2009). Phenomics also offers various platforms connected with computational systems to analyze the phenotyping data including support vector machines (SVMs), artificial neural networks (ANNs), and principal component analysis (Karkee et al., 2009; Yang et al., 2009). So far, the major challenge in plant phenomics is to organize information systems to store datasets of phenotyping/traits so that they may be reanalyzed to generate new ideas (Cabrera-Bosquet et al., 2012).

The combination of GWAS and a high through-put rice phenotyping facility (HRPF) has resulted in the identification of 15 agronomic traits and 25 associated loci corresponding to the Green Revolution semi-dwarf gene (SD1) in rice (Yang et al., 2014). The multifunctional phenotyping technique was based on rice automatic phenotyping (RAP) and the yield traits scorer (YTS), which paved the way for high throughput phenotyping (HTP) methods to replace traditional phenotyping (manual phenotyping) approaches in crop breeding (Yang et al., 2014). Recently, an ontology-driven phenotyping hybrid information system (PHIS) has been proposed to assemble and share multi-scale data and metadata (Neveu et al., 2019). The ontology-driven PHIS information system is a powerful tool to integrate, manage data and share multi-source/multi-scale data (for both greenhouse and field condition), however, ontology-driven architecture creates relationships between objects and enriches datasets with knowledge and metadata (Neveu et al., 2019). Current research efforts have also presented the Internet of Things (IoT)-based CropSight system, which is used to scale and determine both crop phenotyping and genotype–environment interactions (GxE). Internet of Things technologies is worldwide network which uses information and communications technologies for interconnection of sensing and actuating devices providing the ability to share information across platforms. This system can carry out high-quality crop phenotyping and monitor the dynamics of microclimate conditions and has been applied in field wheat crop experiments (Reynolds et al., 2019). However, Roitsch et al. (2019) proposed HTP applications with new generation sensors for next generation phenomics that would contribute to improving crop yields, stress tolerance, and management in the near future.

Overall, phenomics plays an important role in the development of crop breeding strategies through the integration of phenomics with other omics, such as genomics, proteomics, and metabolomics, to provide insights regarding the complex interactions between phenomes, the genome, and environmental factors, which will be beneficial for improving crop management.

ROLE OF BIOINFORMATICS WITHIN THE CONTEXT OF DATABASES AND SOFTWARE TOOLS FOR CROP OMICS ANALYSIS

Bioinformatics is an application of computational technology to handle and analyze the biological data. Bioinformatics as an interdisciplinary field encompasses computer science,

statistics, mathematics, engineering, molecular biology, and biotechnology. Bioinformatics helps in order to interpret biological queries utilizing computational software (Raza et al., 2021). Notably, the integration of omics approaches (i.e., genomics, transcriptomics, proteomics, and metabolomics) has increased our understanding molecular processes associated with abiotic stress responses in plants (Cramer et al., 2011; Jogaiah et al., 2013). Nonetheless, bioinformatics consolidates with these omics approaches and provides base for collecting information for plant abiotic stresses (Ambrosino et al., 2020). Thus, bioinformatics is indispensable for data mining and organization (data production) in support of different omics technologies (Ambrosino et al., 2020). Furthermore, bioinformatics interprets information about the functional system of genes provided by such robust technologies. Bioinformatics also provides accessible resources for computational modeling and simulation analysis by integrating multiple omics technologies. The bioinformatics tools utilizing various software packages have been used for analyzing the multi-omics approaches in crop science. Recently, the availability and advancements of omics platforms have expanded remarkably to allow information to be utilized in multi-dimensional research in the plant sciences. The computational resources have not only made it possible to store, catalog, and analyze the available data but have also provided an easy means to access user friendly databases. Various multi-omics databases have been developed for the crop sciences (Table 1).

Among these, Gramene, Plant Reactome, GabiPD, KaPPA-View4 KEGG, and PMND provide multi-omics-based integration of genomics, transcriptomics, proteomics, and metabolomics for several crop species. Both KNApSACK and KOMICS are useful metabolomics databases that provide information on abundant metabolites in medical plants and crop species. The KOMICS contains several databases for metabolome analysis for example Food Metabolome Repository that can be used to obtain data from various Japanese foods using liquid chromatography-mass spectrometry (LC-MS), KomicMarket database which is used for detected peaks (known/unknown) in metabolome analysis. Similarly Metabolonote database available at KOMICS can also be used to manage “metadata” for experimental 616 data obtained through the metabolomics studies.

PlantTFDB is a multi-crop species database that predicts plant TFs. Moreover, single-crop species dedicated databases of important crops are also available, such as RAP-DB (rice), TFGD (tomato), SoyKB (soybean), MaizeGDB (maize), CerealsDB (wheat), RiceXPro (rice), and SiFGD (Foxtail millet; Table 1). These databases provide comprehensive data on functional genomics coupled with transcriptomics, proteomics, and metabolomics and are currently playing pivotal roles in breeding sciences. Furthermore, there are more than 50 databases numerous crops that provide accumulated omics analysis data (for detailed information see Table 1).

Various software packages have also been developed for multi-omics analysis. In this regard, various online tools have been compiled and are presented in Table 2. The software packages are important to analyze the phenotyping, measurement, and disease symptoms of leaf such as, BioLeaf and EasyPCC. Whereas,

TABLE 1 | List of online databases used for crop multi-omics analysis.

S. no.	Database	Crop species	Features and functionality	Availability/URL
1	Gramene	Multi-species	Multi-omics: comparative functional genomics, transcriptomics, and metabolic pathways	http://www.gramene.org/
2	Plant Reactome	Multi-species	Multi-omics: genomics, transcriptomics, functional proteomics, and integrated metabolic pathways	http://plants.reactome.org/
3	GabiPD	Multi-species	Multi-omics	http://www.gabipd.org/
4	KaPPA-View4 KEGG	Multi-species	Multi-omics: Metabolome integrated transcriptomic and genomic pathway	http://kpv.kazusa.or.jp/kpv4/kegg
5	KOMICS	Multi-species	Metabolomics	http://www.kazusa.or.jp/komics/en/
6	KNAPSAck	Multi-species	Metabolomics	http://kanaya.naist.jp/KNAPSAck/Family/
7	PMND	Multi-species	Multi-omics: Metabolome integrated transcriptome and genomic pathway	https://www.plantcyc.org/
8	PlantTFDB	Multi-species	Transcriptomics: Predicts plant transcription factors (TFs)	http://planttfdb.cbi.pku.edu.cn/
9	PPDB	Multi-species	Proteomics	http://ppdb.tc.cornell.edu/
10	GrainGenes	Multi-species	Genomics	http://www.graingenes.org
11	PlantGDB	Multi-species	Comparative genomics	http://www.plantgdb.org/
12	PCD	Multi-species	Genomics-assisted breeding (GAB)	https://www.pulsedb.org/
13	RAP-DB	Rice	Genomics integrated multi-omics	http://rapdb.dna.affrc.go.jp/
14	RiceXPro	Rice	Functional genomics and transcriptomics	http://ricexpro.dna.affrc.go.jp/
15	Ricebase	Rice	Genomics	https://ricebase.org/
16	Oryzabase	Rice	Integrated biological and genome information database	https://shigen.nig.ac.jp/rice/oryzabase/
17	SNP-Seek II	Rice	SNP-seek database	http://snp/seek.irri.org
18	RicyerDB	Rice	Integrative genomics and proteomics	http://server.malab.cn/Ricyer/index.html
19	RiceVarMap	Rice	Genomic variation and functional annotation	http://ricevarmap.ncpgr.cn
20	TFGD	Tomato	Functional genomics, integrated transcriptomics, and metabolomics	http://ted.bti.cornell.edu/
21	TOMATOMICS	Tomato	Multi-omics	http://plantomics.mind.meiji.ac.jp/tomatomics/index.html
22	SoyKB	Soybean	Multi-omics	http://soykb.org/
23	SoyBase	Soybean	Multi-omics	https://soybase.org/
24	SoyGD	Soybean	Genomics	http://soybeangenome.siu.edu/
25	SFGD	Soybean	Multi-omics: Functional genomics, transcriptomics, and metabolic pathways	http://bioinformatics.cau.edu.cn/SFGD/
26	SoyNet	Soybean	Functional genomics and transcriptomics	https://www.inetbio.org/soynet/
27	MaizeGDB	Maize	Multi-omics	https://www.maizegdb.org/
28	MaizeDIG	Maize	Phenomics and genomics	https://maizedig.maizegdb.org/
29	MaizeSNPDB	Maize	SNPs	https://github.com/venyao/MaizeSNPDB
30	MMAD	Maize	Microarray	http://maizearrayannot.bi.up.ac.za/
31	CSRDB	Maize and rice	Small RNAs database	http://sundarlab.ucdavis.edu/smrnas/
32	RGPDB	Maize, soybean and sorghum	Multi-omics	http://sysbio.unl.edu/RGPDB/
33	CerealsDB	Wheat	Functional genomics	http://www.cerealsdb.uk.net/cerealgeno/mics/
34	WGI	Wheat	Genomics	http://wheatgenome.info/
35	wDBTF	Wheat	Transcription factors	http://www.appli.nantes.inra.fr:8180/wDBTF/

(Continued)

TABLE 1 | Continued

S. no.	Database	Crop species	Features and functionality	Availability/URL
36	WheatGPE	Wheat	Base on phenotype-genotype and environment	http://www.wheatdb.org/
37	SIFGD	Foxtail millet	Functional genomics integrated transcriptomics, and metabolomics	http://structuralbiology.cau.edu.cn/SIFGD/
38	CottonFGD	Cotton	Functional genomics	https://cottonfgd.org/
39	CottonGen	Cotton	Genomics	https://www.cottongen.org/
40	CottonQTLdb	Cotton	QTL analysis	http://www.cottonqtlidb.org
41	MTGD	<i>Medicago truncatula</i>	Genomics	http://www.MedicagoGenome.org
42	CTDB	Chickpea	Functional genomics and transcriptomics	http://www.nipgr.ac.in/ctdb.html
43	ECPD	Potato	Genomics	https://www.europotato.org/
44	BARLEX	Barley	Genomics	http://barlex.barleysequence.org
45	SorghumFDB	Sorghum	Functional genomics	http://structuralbiology.cau.edu.cn/sorghum/index.html
46	SorGSD	Sorghum	SNPs	http://sorgsd.big.ac.cn/
47	SGH	Sugarcane	Genomics	https://sugarcane-genome.cirad.fr/
48	BRAD	<i>Brassica</i> species	Genomics and transcriptomics	http://brassicadb.org/brad/
49	CropSNPdb	<i>Brassica</i> species and wheat	SNPs	http://snpdb.appliedbioinformatics.com.au/
50	SFGD	Sunflower	Genomics and transcriptomics	https://www.sunflowergenome.org
51	BBDG	Blueberry	Genomics and transcriptomics	http://bioinformatics.towson.edu/BBDG/

GABI, genomanalyse im biologischen system pflanze; GABI, GABI primary database; PMND, plant metabolic network database; SIFGD, *Setaria italica* functional genomics database; RAP-DB, the rice annotation project database; TFGD, tomato functional genomics database; TFDB, transcription factors database; PPDB, plant proteome database; PCD, pulse crop database; RiceVarMap, rice variation map; SoyGD, soybean genome database; SFGD, soybean functional genomics database; SoyNet, soybean network; MaizeGDB, maize genetics and genomics database; MaizeDIG, database of images and genomes; MMAD, maize microarray annotation database; CSRDB, cereal small RNAs database; RGPDB, root genes and promoters database; WGI, wheat genome info; wDBTF, wheat database for transcription factor; WheatGPE, Genotype-phenotype and environment; CottonFGD, CottonFGD cotton functional genomics database; CottonGen, cotton genomics; MtGD, *Medicago truncatula* genome database; CTDB, chickpea transcriptome database; ECPD, European cultivated potato database; SorghumFDB, Sorghum functional genomics database; SorGSD, Sorghum genome SNP database; SGH, sugarcane genome hub; BRAD, (*Brassica* database; SFGD, sunflower genome database; BBDG, blueberry genomics database.

protein–protein interaction, gene structure analysis can also predicted in STRING and GSDS, respectively. However, Gromacs software could be used for simulation of protein and lipids for crops (for detail see Table 2).

Several softwares can be used as individual omics analysis such as, phenomics, proteomics, and metabolomics, whereas, some of them are useful for multi-omics analysis. In detail, LemnaLauncher, BioLeaf, and EasyPCC are used for phenomics analysis of crops. For proteome function and interactions, three important software packages (i.e., STRING, SPPS, and ProteinProspector) provide vital information. Importantly, Gromacs and PTools are multi-omics based software for multiple crop species. In order to analyze structural and comparative genomics, GSDS, GAP4, and VISTA are accessible tools for crop omics. Additionally, AMDIS and SIMCA-P 14.0 could be used for ionome-integrated metabolic component analysis in crop species (Table 2).

ROLE OF PANOMICS FOR CROP BREEDING SCIENCE

Panomics provides a platform to integrate complex omics, such as genomics, epigenomics, transcriptomics, proteomics, PTM proteomics, metabolomics, and phenomics (Weckwerth et al., 2020). The concept of panomics was recently proposed

by Weckwerth et al. (2020). The idea of this platform is to combine different omics and construct models that can be used to predict complex traits (Weckwerth, 2011a, 2019). However, coupling phenomics and environmental information with genomics, transcriptomics, proteomics, and metabolomics would provide a better understanding of the terroir-phenotype dependency at a molecular level. The integration of complex “omics” datasets could also reduce the number of false positives generated from single data sources for genotype-phenotype prediction (Ritchie et al., 2015). Panomics and environmental platforms together with multiple data integration can be used to identify genes, QTLs, and markers through functional omics and mathematical models to enhance the tolerance to abiotic and biotic stress in crop varieties and create elite lines to improve the germplasm (Weckwerth et al., 2020). To analyze the integrated data, special tools can be used to merge multi-omics datasets prior to any interpretation (Kuo et al., 2013) (e.g., tools such as PAINTOMICS, KaPPA-view, and COVAIN). PAINTOMICS is web based tool offers integrated visualization of data of two omics, the transcriptomics and metabolomics datasets and displays the data on KEGG pathway maps (García-Alcalde et al., 2011). Another web based tool the KaPPA-view has been developed for integration of transcript and metabolite data on plant metabolic pathway maps (Tokimatsu et al., 2005). Nevertheless, the COVAIN (covariance inverse) tool primarily

TABLE 2 | List of online software packages used for crop multi-omics analysis.

S. no.	Software	Crop species	Features and functionality	Availability/URL
1	LemnaLauncher	Multi-species	Phenomics: image-based measurements of length, width, color, surface of seeds	http://www.plant/image/analysis.org/software/lemnalauncher
2	EasyPCC	Multi-species	Phenomics: field crop canopy measurement	http://www.plant/image/analysis.org/software/easypcc
3	BioLeaf	Multi-species	Phenomics: leaf surface and disease analysis	http://www.plant/image/analysis.org/software/bioleaf
4	STRING	Multi-species	Proteomics: predicts protein interactions containing functional associations	http://string.embl.de
5	SPPS	Multi-species	Proteomics: predicts protein–protein interaction partners	http://mdl.shsmu.edu.cn/SPPS/
6	ProteinProspector	Multi-species	Proteomics: sequence mining with MS	http://prospector.ucsf.edu/
7	Gromacs	Multiple-species	Genomics, proteomics and metabolomics	https://omictools.com/gromacs/tool
8	PTools	Multi-species	Multi-omics	https://omictools.com/ptools/tool
9	GSDS	Multi-species	Structural genomics: visualizes gene structure (exons, introns, and UTRs).	http://gsds.cbi.pku.edu.cn/
10	GAP4	Multi-species	Structural genomic sequence assembly	http://stadensourceforge.net/overview.html
11	VISTA	Multi-species	Comparative genomics	http://genome.lbl.gov/vista/index.shtml
12	AMDIS	Multi-species	Metabolomics: GC-MS data interpretation	http://www.amdis.net/
13	SIMCA-P 14.0	Multi-species	Ionomics integrated with metabolome: principal metabolic component analysis	https://umetrics.com/kb/simca/online/140

SPPS, sequence-based protein partners search, GAP4, genome assembly program; GSDS, gene structure display server.

used for metabolomics data and can support in statistical analysis of the integrated omics dataset with KEGG pathway and gene ontology analysis (Sun and Weckwerth, 2012). The integration of GWAS with panomics has also been used to explain and understand phenotypic variance in crops. Importantly, integrating GWAS with omics datasets including transcriptomics (eQTLs), proteomics (pQTLs), and metabolomics (mQTLs) may lead to the identification of novel genes and functional pathways underlying complex traits (Weckwerth et al., 2020). In this vein, a combined metabolome-based genome-wide association study (mGWAS with eQTL) identified metabolite features associated with kernel weight in maize crop (Wen et al., 2014).

Furthermore, the integration of panomics and genome editing tools (e.g., TALENs and CRISPR/Cas9) has been proposed as a model for the development of precision breeding (Weckwerth et al., 2020). Recently, using MAS and genomic selection techniques, agronomically important genes have been identified that only explain ~40% of the phenotypic variance. Hence, the proposed methodology of the integration of panomics with genome editing tools could result in the identification of the remaining ~60% of the phenotypic variance and may support the identification of agronomically important genes in a fast and effective manner to support precision breeding efforts. Thus, this methodology will not only be helpful for improving crops but will also ensure precision in trait optimization in

terms of yield, nutritional value, and plant fitness (Weckwerth et al., 2020). Hence, genotype to phenotype concept based on epigenetic regulation (triggered by environmental factors) through integration of multi-omics could lead to develop qualitative and quantitative traits which may be helpful for crop breeding improvement (Figure 2).

INTEGRATION OF MULTI-OMICS AND SYSTEMS BIOLOGY APPROACHES FOR CROP IMPROVEMENT

In order to understand the cellular components and complex behaviors of biological systems, an integration of the different omics approaches is required to envisage the responses of a given organism under a set of conditions. Previously, the coupling of metabolomics with genomics, transcriptomics, and proteomics provided an integrated portrait of functions, spanning the genome to phenotypic interactions with the environment (Weckwerth, 2011a). Combined omics approaches have been applied in potato tubers and Arabidopsis to analyze transcriptomic and metabolomic profiles (Urbanczyk-Wochniak et al., 2003; Hirai et al., 2004). These studies have demonstrated that coupling of different omics approaches could be useful for identifying potential candidate genes for functional

analysis. Since the advancements in omics technologies and computational tools, integrative omics approaches have been implemented in the crop sciences. For example, the epigenetic-based integration of multi-omics revealed the role of the regulation of lipid biosynthesis during cotton fiber development (Wang et al., 2016). The integration of GWAS with metabolite profiling strategies has proved to be a powerful technique to dissect the biochemical and genetic processes in several model crop species including rice, maize, and tomato (Luo, 2015; Matsuda et al., 2015). Importantly, the integration of omics approaches (i.e., genomics, transcriptomics, proteomics, and metabolomics) has led to abiotic stress tolerant crop phenotypes (Jogaiah et al., 2013). Functional genomics and mutagenomics have been used to identify numerous mutants with specific variations with regard to growth, development, and stress tolerance in various crops including rice, maize, wheat, and barley (Talukdar and Sinjushin, 2015). Combined GWAS and HRP approach was able to elucidate agronomic traits responsible for biomass growth and yield in rice crop (Yang et al., 2014). This robust technique replaced the traditional phenomics, providing a powerful tool for crop genetics and breeding sciences (Yang et al., 2014). Combined GWAS and high-throughput leaf scoring (HLS) was used to identify new loci related to the size, shape, and color of leaves in rice crop (Yang et al., 2015). The performance of QTL mapping combined with agronomic traits also helped to identify numerous QTLs in maize crop (Zhang et al., 2017). Hence, genomic information combined with potential phenotyping approaches can provide information on complex traits to improve crops (Zhao et al., 2019). Combined omics approaches could complement each other when analyzing certain biological processes. This idea has been validated through the differential regulation of metabolites, proteins, and ions related to salinity stress in halophytes (Kumari et al., 2015). Metabolomics is considered to be a link between genotypes and phenotypes (Fiehn, 2002). Combined ionome and metabolome techniques were used to suppress photosynthesis and growth rates in maize crop under alkaline conditions (Guo et al., 2017). Similarly, the leaf and grain ionome revealed mineral element genetic diversification in rice crop through genetic mapping and QTL identification (Norton et al., 2010; Zhang et al., 2014b; Pinson et al., 2015). Thus, genotype to phenotype-based integration of multi-omics would provide insights into the functional mechanisms of genes and their networks to improve crop science, genetics, growth, yield, and resistance in response to physiological and environmental stress (Figure 2).

Systems biology attempts to understand the complete biological system through modeling. It predicts the behavior of all components and interactions among genes, proteins, and metabolites with respect to external stimuli (Kumar et al., 2015a). Systems biology has provided a powerful base to combine multi-omics to create a holistic understanding of an organism related to its adaptation and development (Pinu et al., 2019). Multi-omics approaches have been employed in plant stress research associated with systems biology (Mosa et al., 2017). However, comprehensive analyses using three omics technologies, transcriptomics, metabolomics, and proteomics, have also increased our understanding of systems biology

associated with abiotic stress responses in plants (Cramer et al., 2011). Multi-omics integrated with systems biology based on top-down and bottom-up data reduction approaches, which employ genomics and/or metabolomics as a foundation, is able to predict phenotypic responses and metabolic pathways (Pinu et al., 2019). Another study proposed two system-based approaches for decoding the complexity of biological systems. First, top-down or integrative systems biology has been employed with high-throughput multi-omics data and data analysis using bioinformatics and systems biology tools to identify agriculturally important traits. Second, bottom-up or predictive systems biology in which the properties of genes or proteins with available quantitative information are utilized to develop models of well-characterized components of both genes and proteins has been used to predict the behavior of systems under different conditions (Kumar et al., 2015a). Hence, a model needs to be developed and linked to phenotypic traits to allow for valuable progress with regard to genetic manipulation and crop production. The integration of multi-omics and systems biology approaches has resulted in the identification of molecular regulator networks for salt stress tolerance in grapevine crop (Daldoul et al., 2014). Moreover, systems biology integrated with omics approaches for network and testing models has been proposed for abiotic stress responses in crop plants (Gupta et al., 2013). In this regard, we proposed top-down (phenotype to genotype) and bottom-up (genotype to phenotype) model based on an integration of multi-omics with systems biology in response to environmental stress, which may also be useful to improve crop breeding (Figure 2).

CONCLUSION AND PERSPECTIVE

Multi-omics analysis has played an integral role in the identification of genetic processes, growth, development, and stress tolerance in various crops. Several omics approaches including genomics, transcriptomics, proteomics, metabolomics, ionomics, and phenomics have employed high throughput techniques to interpret functional analysis, molecular mechanisms of genes, and gene networks in crop science. Furthermore, the integration of GWAS with metabolomics, transcriptomics, and proteomics has proved to be a potential tool to elucidate biochemical processes and abiotic stress tolerance in some model crops. The studies have shown that how the combination of several omics approaches could be beneficial for identifying potential candidate genes and their pathways. With advances in high throughput technologies and computational tools, the integration of some omics approaches has been possible in the crop sciences. The panomics platform with integrated multi-omics, such as genomics, epigenomics, transcriptomics, proteomics, metabolomics, and phenomics, would facilitate the construction of models to predict agronomically important traits to improve crops through precision breeding. Importantly, the integration of systems biology with complex omics datasets has also increased our understanding of molecular regulator networks for crop improvement. The studies have revealed the G–P–E interactions

in crops. Subsequently, integration of functional genomics with transcriptomics, proteomics, metabolomics, and ionomics may result in apparent crop quality phenotypic traits under certain stresses through “genotype to phenotype” concept. From this perspective, we propose, the integration of multi-omics with systems biology by top-down (phenotype to genotype) and bottom-up (genotype to phenotype) model that can be helpful to develop quality agronomic traits for crop improvements under environmental stress conditions (Figure 2).

AUTHOR CONTRIBUTIONS

YY, MS, WA, YW, and JZ drafted the manuscript. YY, JZ, JL, and MS collected the background information. MHS and FW

analyzed the databases and software tools for omics. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Natural Science Foundation of China (No. 31870670), the fundamental Scientific Research Funds for Chinese Academy of Tropical Agriculture Sciences (CATAS-No. 1630152017019), Central Public-interest Scientific Institution Basal Research Fund for Innovative Research Team Program of CATAS (NO.17CXTD-28), the earmarked fund for the Belt and Road Tropical Project (BARTP-06), and the Talented Young Scientist Program (TYSP) from Ministry of Science and Technology of China.

REFERENCES

- Abdelrahman, M., El-Sayed, M. A., Hashem, A., Abd-Allah, E. F., Alqarawi, A. A., Burritt, D. J., et al. (2018). Metabolomics and transcriptomics in legumes under phosphate deficiency in relation to nitrogen fixation by root nodules. *Front. Plant Sci.* 9:922. doi: 10.3389/fpls.2018.00922
- Agarwal, R., Bentur, J. S., and Nair, S. (2014). Gas chromatography mass spectrometry based metabolic profiling reveals biomarkers involved in rice-gall midge interactions. *J. Integr. Plant Biol.* 56, 837–848. doi: 10.1111/jipb.12244
- Agarwal, M., Shrivastava, N., and Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep.* 27, 617–631. doi: 10.1007/s00299-008-0507-z
- Ahloowalia, B. S., Maluszynski, M., and Nichterlein, K. (2004). Global impact of mutation derived varieties. *Euphytica* 135, 187–204. doi: 10.1023/B:EUPH.0000014914.85465.4f
- Aizat, W. M., and Hassan, M. (2018). “Proteomics in systems biology,” in *Omics Applications for Systems Biology. Advances in Experimental Medicine and Biology*, eds W. Aizat, H. H. Goh, and S. Baharum (Cham: Springer), 31–49. doi: 10.1007/978-3-319-98758-3_3
- Allwood, J. W., and Goodacre, R. (2010). An introduction to liquid chromatography-mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochem. Anal.* 21, 33–47. doi: 10.1002/pca.1187
- Alseikh, S., Aharoni, A., Brotman, Y., Contrepoint, K., D'Auria, J., Ewald, J., et al. (2021). Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nat. Methods* 18, 747–756. doi: 10.1038/s41592-021-01197-1
- Ambrosino, L., Colantuono, C., Diretto, G., Fiore, A., and Chiusano, M. L. (2020). Bioinformatics resources for plant abiotic stress responses: state of the art and opportunities in the fast evolving-omics era. *Plants* 9:591. doi: 10.3390/plants9050591
- Appleby, N., Edwards, D., and Batley, J. (2009). “New technologies for ultra-high throughput genotyping in plants,” in *Methods in Molecular Biology, Plant Genomics*, eds J. P. Gustafson, P. Langridge, D. J. Somers, and N. J. Totowa (New York, NY: Humana Press), 19–39. doi: 10.1007/978-1-59745-427-8_2
- Arefian, M., Vessal, S., Malekzadeh-Shafaroudi, S., Siddique, K. H. M., and Bagheri, A. (2019). Comparative proteomics and gene expression analyses revealed responsive proteins and mechanisms for salt tolerance in chickpea genotypes. *BMC Plant Biol.* 19:300. doi: 10.1186/s12870-019-1793-z
- Aslam, B., Basit, M., Atif, N. M., Khurshid, M., and Rasool, M. H. (2017). Proteomics: technologies and their applications. *J. Chromato Sci.* 55, 182–196. doi: 10.1093/chromsci/bmw167
- Baggerman, G., Geert, Vierstraete, E., De, L. A., and Schoofs, L. (2005). Gel-based versus gel-free proteomics: a review. *Comb. Chem. High Throughput Screen.* 8, 669–677. doi: 10.2174/138620705774962490
- Baharum, S. N., and Azizan, K. A. (2018). Metabolomics in systems biology. *Adv. Exp. Med. Biol.* 1102, 51–68. doi: 10.1007/978-3-319-98758-3_4
- Baker, N. R. (2008). Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol.* 59, 89–113. doi: 10.1146/annurev.arplant.59.032607.092759
- Balmerl, D., Flors, V., Glauser, G., and Mauch-Mani, B. (2013). Metabolomics of cereals under biotic stress: current knowledge and techniques. *Front. Plant Sci.* 4:82. doi: 10.3389/fpls.2013.00082
- Basha, S. M., Mazhar, H., and Vasanthaiah, H. K. N. (2010). Proteomics approach to identify unique xylem sap proteins in Pierce's disease-tolerant *Vitis* species. *Appl. Biochem. Biotechnol.* 160, 932–944. doi: 10.1007/s12010-009-8620-1
- Bashir, K., Matsui, A., Rasheed, S., and Seki, M. (2019). Recent advances in the characterization of plant transcriptomes in response to drought, salinity, heat, and cold stress. *Fl000Res.* 8:658. doi: 10.12688/f1000research.18424.1
- Baxter, I. (2010). Ionomics: the functional genomics of elements. *Brief. Funct. Genomics* 9, 149–156. doi: 10.1093/bfpg/elp055
- Baxter, I. (2015). Should we treat the ionome as a combination of individual elements, or should we be deriving novel combined traits? *J. Exp. Bot.* 66, 2127–2131. doi: 10.1093/jxb/erv040
- Baxter, I. R., Ziegler, G., Lahner, B., Mickelbart, M. V., Foley, R., Danku, J., et al. (2014). Single-kernel ionomic profiles are highly heritable indicators of genetic and environmental influences on elemental accumulation in maize grain (*Zea mays*). *PLoS ONE* 9:e87628. doi: 10.1371/journal.pone.0087628
- Berger, B., Parent, B., and Tester, M. (2010). High-throughput shoot imaging to study drought responses. *J. Exp. Bot.* 61, 3519–3528. doi: 10.1093/jxb/erq201
- Bhardwaj, A. R., Joshi, G., Kukreja, B., Malik, V., Arora, P., Pandey, R., et al. (2015). Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop *Brassica juncea*. *BMC Plant Biol.* 15:9. doi: 10.1186/s12870-014-0405-1
- Bodner, G., Alsalem, M., Nakhforoosh, A., Arnold, T., and Leitner, D. (2017). RGB and spectral root imaging for plant phenotyping and physiological research: experimental setup and imaging protocols. *J. Vis. Exp.* 126:56251. doi: 10.3791/56251
- Bonthala, V. S., Mayes, K., Moreton, J., Blythe, M., Wright, V., May, S. T., et al. (2016). Identification of gene modules associated with low temperatures response in Bambara groundnut by network-based analysis. *PLoS ONE* 11:e0148771. doi: 10.1371/journal.pone.0148771
- Bowne, J. B., Erwin, T. A., Juttner, J., Schnurbusch, T., Langridge, P., Bacic, A., et al. (2012). Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Mol. Plant* 5, 418–429. doi: 10.1093/mp/ssr114
- Brunetti, C., George, R. M., Tattini, M., Field, K., and Davey, M. P. (2013). Metabolomics in plant environmental physiology. *J. Exp. Bot.* 64, 4011–4020. doi: 10.1093/jxb/ert244
- Burgess, D. J. (2015). Putting transcriptomics in its place. *Nat. Rev. Genet.* 16:319. doi: 10.1038/nrg3951
- Butelli, E., Titta, L., Giorgio, M., Mock, H. P., Matros, A., Peterek, S., et al. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* 26, 1301–1308. doi: 10.1038/nbt.1506

- Cabrera-Bosquet, L., Crossa, J., von Zitzewitz, J., Serret, M. D., and Araus, J. L. (2012). High-throughput phenotyping and genomic selection: the frontiers of crop breeding converge. *J. Integr. Plant. Biol.* 54, 312–320. doi: 10.1111/j.1744-7909.2012.01116.x
- Caldwell, D. G., McCallum, N., Shaw, P., Muehlbauer, G. J., Marshall, D. F., and Waugh, R. (2004). A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). *Plant J.* 40, 143–150. doi: 10.1111/j.1365-313X.2004.02190.x
- Callinan, P. A., and Feinberg, A. P. (2006). The emerging science of epigenomics. *Human Mol. Genet.* 15 (Suppl_1), R95–R101. doi: 10.1093/hmg/ddl095
- Campbell, B. C., Gilding, E. K., Mace, E. S., Tai, S., Tao, Y., Prentis, P. J., et al. (2016). Domestication and the storage starch biosynthesis pathway: signatures of selection from a whole sorghum genome sequencing strategy. *Plant Biotechnol. J.* 14, 2240–2253. doi: 10.1111/pbi.12578
- Cao, A. H., Xing, L. P., Wang, X. Y., Yang, X. M., Wang, W., YSun, Y. L., et al. (2011). Serine/threonine kinase gene Stpk-V, a key member of powdery mildew resistance gene Pm21, confers powdery mildew resistance in wheat. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7727–7732. doi: 10.1073/pnas.1016981108
- Chaerle, L., Lenk, S., Leinonen, I., Jones, H. G., Van Der Straeten, D., and Buschmann, C. (2009). Multi-sensor plant imaging: towards the development of a stress-catalogue. *Biotechnol. J.* 4, 1152–1167. doi: 10.1002/biot.200800242
- Challa, S., and Neelapu, N. R. (2018). “Genome-wide association studies (GWAS) for abiotic stress tolerance in plants,” in *Biochemical, Physiological and Molecular Avenues for Combating Abiotic Stress Tolerance in Plants*, eds S. H. Shabir Hussain Wani (Cham: Academic Press), 135–150. doi: 10.1016/B978-0-12-813066-7.00009-7
- Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. *Curr. Opin. Plant Biol.* 27, 207–216. doi: 10.1016/j.pbi.2015.08.003
- Chen, H., Wang, L., Liu, X., Hu, L., Wang, S., and Cheng, X. (2017). *De novo* transcriptomic analysis of cowpea (*Vigna unguiculata* L. Walp.) for genic SSR marker development. *BMC Genet.* 18:65. doi: 10.1186/s12863-017-0531-5
- Chen, L., Hao, L., Parry, M. A. J., Phillips, A. L., and Hu, Y. G. (2014). Progress in TILLING as a tool for functional genomics and improvement of crops. *J. Integr. Plant Biol.* 56, 425–443. doi: 10.1111/jipb.12192
- Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., et al. (2014). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452, 215–219. doi: 10.1038/nature06745
- Cooper, J. L., Till, B. J., Laport, R. G., Darlow, M. C., Kleffner, J. M., Jamai, A., et al. (2008). TILLING to detect induced mutations in soybean. *BMC Plant Biol.* 8:9. doi: 10.1186/1471-2229-8-9
- Cramer, G. R., Urano, K., Delrot, S., Pezzotti, M., and Shinozaki, K. (2011). Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biol.* 11:163. doi: 10.1186/1471-2229-11-163
- Daldoul, S., Ben, Amar, A., Guillaumie, S., and Mliki, A. (2014). Integration of omics and system biology approaches to study grapevine (*Vitis vinifera* L.) response to salt stress: a perspective for functional genomics - a review. *OENO One* 48, 189–200. doi: 10.20870/oeno-one.2014.48.3.1573
- De Cremer, K., Mathys, J., Vos, C., Froeniche, L., Michelmore, R. W., Cammue, B. P. A., et al. (2013). RNAseq-based transcriptome analysis of *Lactuca sativa* infected by the fungal necrotroph *Botrytis cinerea*. *Plant Cell Environ.* 36, 1992–2007. doi: 10.1111/pce.12106
- Deeba, F., Pandey, A. K., Ranjan, S., Mishra, A., Singh, R., Sharma, Y. K., et al. (2012). Physiological and proteomic responses of cotton (*Gossypium herbaceum* L.) to drought stress. *Plant Physiol. Biochem.* 53, 6–18. doi: 10.1016/j.plaphy.2012.01.002
- Demirevska, K., Simova-Stoilova, L., Vassileva, V., Vaseva, I., Grigorova, B., and Feller, U. (2008). Drought-induced leaf protein alterations in sensitive and tolerant wheat varieties. *Gen. Appl. Plant Physiol.* 34, 79–102. doi: 10.7892/boris.110728
- Deshmukh, R., Sonah, H., Patil, G., Chen, W., Prince, S., Mutava, R., et al. (2014). Integrating omic approaches for abiotic stress tolerance in soybean. *Front. Plant Sci.* 25:244. doi: 10.3389/fpls.2014.00244
- Di-Baccio, D., Navari-Izzo, F., and Izzo, R. (2004). Seawater irrigation: antioxidant defence responses in leaves and roots of a sunflower (*Helianthus annuus* L.) ecotype. *J. Plant Physiol.* 161, 1359–1366. doi: 10.1016/j.jplph.2003.07.001
- Ding, X., Li, X., and Xiong, L. (2013). Insight into differential responses of upland and paddy rice to drought stress by comparative expression profiling analysis. *Int. J. Mol. Sci.* 14, 5214–5238. doi: 10.3390/ijms14035214
- Do, P. T., Degenkolbe, T., Erban, A., Heyer, A. G., Kopka, J., Kohl, K. I., et al. (2013). Dissecting rice polyamine metabolism under controlled long-term drought stress. *PLoS ONE* 8:e60325. doi: 10.1371/journal.pone.0060325
- Dong, C., Dalton-Morgan, J., Vincent, K., and Sharp, P. (2009). A modified tillage method for wheat breeding. *Plant Gen.* 2, 39–47. doi: 10.3835/plantgenome2008.10.0012
- Dugas, D. V., Monaco, M. K., Olsen, A., Klein, R. R., Kumari, S., Ware, D., et al. (2011). Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid. *BMC Genomics* 12:514. doi: 10.1186/1471-2164-12-514
- Duque, A. S., Almeida, A. M., Bernardes, da Silva, A., Marques, da Silva, J., Farinha, A. P., Santos, D., et al. (2013). “Chapter 3: Abiotic stress responses in plants: unraveling the complexity of genes and networks to survive,” in *Abiotic Stress: Plant Responses and Applications in Agriculture*, eds K. Vahdati and C. Leslie (Rijeka: INTECH Open), 49–102. doi: 10.5772/45842
- Dwivedi, S., Perotti, E., and Ortiz, R. (2008). Towards molecular breeding of reproductive traits in cereal crops. *Plant Biotechnol. J.* 6, 529–559. doi: 10.1111/j.1467-7652.2008.00343.x
- Eldakak, M., Milad, S. I., Nawar, A. I., and Rohila, J. S. (2013). Proteomics: a biotechnology tool for crop improvement. *Front. Plant Sci.* 4:35. doi: 10.3389/fpls.2013.00035
- El-Metwally, S., Ouda, O. M., and Helmy, M. (2014). *Next Generation Sequencing Technologies and Challenges in Sequence Assembly, 1st Edn.* New York, NY: Springer. doi: 10.1007/978-1-4939-0715-1
- Fiehn, O. (2002). Metabolomics — the link between genotypes and phenotypes. *Plant Mol. Biol.* 48, 155–171. doi: 10.1023/A:1013713905833
- Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* 18, 1157–1161. doi: 10.1038/81137
- Finkel, E. (2009). With ‘phenomics’ plant scientists hope to shift breeding into overdrive. *Science* 325, 380–381. doi: 10.1126/science.325_380
- Fournier, M. L., Gilmore, J. M., Martin-Brown, S. A., and Washburn, M. P. (2007). Multidimensional separations-based shotgun proteomics. *Chem. Rev.* 107, 3654–3686. doi: 10.1021/cr068279a
- Francia, E., Morcia, C., Pasquariello, M., Mazzamurro, V., Milc, J. A., Rizza, F., et al. (2016). Copy number variation at the HvCBF4-HvCBF2 genomic segment is a major component of frost resistance in barley. *Plant Mol. Biol.* 92, 161–175. doi: 10.1007/s11103-016-0505-4
- Fukuoka, S., Saka, N., Koga, H., Ono, K., Shimizu, T., Ebana, K., et al. (2009). Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science* 325, 998–1001. doi: 10.1126/science.1175550
- Furbank, R. T., and Tester, M. (2011). Phenomics—technologies to relieve the phenotyping bottleneck. *Tred. Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005
- Gabur, I., Chawla, H. S., Liu, X., Kumar, V., Faure, S., von Tiedemann, S., et al. (2018). Finding invisible quantitative trait loci with missing data. *Plant Biotechnol. J.* 16, 2102–2112. doi: 10.1111/pbi.12942
- García-Alcalde, F., García-López, F., Dopazo, J., and Conesa, A. (2011). Paintomics: a web based tool for the joint visualization of transcriptomics and metabolomics data. *Bioinformatics* 27, 137–139. doi: 10.1093/bioinformatics/btq594
- Gent, J. I., Ellis, N. A., Guo, L., Harkess, A. E., Yao, Y., Zhang, X., et al. (2013). CHH islands: *de novo* DNA methylation in near-gene chromatin regulation in maize. *Genome Res.* 23, 628–637. doi: 10.1101/gr.146985.112
- Ghatak, A., Chaturvedi, P., and Weckwerth, W. (2017). Cereal crop proteomics: systemic analysis of crop drought stress responses towards marker-assisted selection breeding. *Front. Plant Sci.* 8:757. doi: 10.3389/fpls.2017.00757
- Ghatak, A., Chaturvedi, P., and Weckwerth, W. (2018). Metabolomics in plant stress physiology. *Adv. Biochem. Eng. Biotechnol.* 164:187–236. doi: 10.1007/10_2017_55
- Ghosh, D., and Xu, J. (2014). Abiotic stress responses in plant roots: a proteomics perspective. *Front. Plant Sci.* 5:6. doi: 10.3389/fpls.2014.0006
- Gjivusland, A. B., Vik, J. O., Beard, D. A., Hunter, P. J., and Omholt, S. W. (2013). Bridging the genotype-phenotype gap: what does it take? *J. Physiol.* 591, 2055–2066. doi: 10.1113/jphysiol.2012.248864
- Goh, H. H. (2018). “Integrative multi-omics through bioinformatics,” in *Omics Applications for Systems Biology. Advances in Experimental Medicine and*

- Biology, eds W. Aizat, H. H. Goh, and S. Baharum (Cham: Springer), 69–80. doi: 10.1007/978-3-319-98758-3_5
- Golicz, A. A., Batley, J., and Edwards, D. (2016a). Towards plant pangenomics. *Plant Biotechnol. J.* 14, 1099–1105. doi: 10.1111/pbi.12499
- Golicz, A. A., Bayer, P. E., Barker, G. C., Edger, P. P., Kim, H., Martinez, P. A., et al. (2016b). The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nat. Commun.* 7:13390. doi: 10.1038/ncomms13390
- González, R. M., Ricardi, M. M., and Iusem, N. D. (2013). Epigenetic marks in an adaptive water stress responsive gene in tomato roots under normal and drought conditions. *Epigenetics* 8, 864–872. doi: 10.4161/epi.25524
- Gordon, S. P., Contreras-Moreira, B., Woods, D. P., Marais, D. L. D., Burgess, D., Shu, S. Q., et al. (2017). Extensive gene content variation in the *Brachypodium distachyon* pan-genome correlates with population structure. *Nat. Commun.* 8:2184. doi: 10.1038/s41467-017-02292-8
- Großkinsky, D. K., Syaifullah, S. J., and Roitsch, T. (2018). Integration of multi-omics techniques and physiological phenotyping within a holistic phenomics approach to study senescence in model and crop plants. *J. Exp. Bot.* 69, 825–844. doi: 10.1093/jxb/erx333
- Gu, R., Chen, F., Liu, B., Wang, X., Liu, J., Li, P., et al. (2015). Comprehensive phenotypic analysis and quantitative trait locus identification for grain mineral concentration, content, and yield in maize (*Zea mays* L.). *Theor. Appl. Genet.* 128, 1777–1789. doi: 10.1007/s00122-015-2546-5
- Gunnaiah, R., Kushalappa, A. C., Duggavathi, R., Fox, S., and Somers, D. J. (2012). Integrated metabolomic approach to decipher the mechanisms by which wheat QTL (Fhb1) contributes to resistance against *Fusarium graminearum*. *PLoS ONE* 7:e40695. doi: 10.1371/journal.pone.0040695
- Guo, P., Baum, M., Grando, S., Ceccarelli, S., Bai, G., Li, R. V. K., et al. (2009). Differentially expressed genes between drought-tolerant and drought-sensitive barley genotypes in response to drought stress during the reproductive stage. *J. Exp. Bot.* 60, 3531–3544. doi: 10.1093/jxb/erp194
- Guo, R., Shi, L., Yan, C., Zhong, X., Gu, F., Liu, Q., et al. (2017). Ionic and metabolic responses to neutral salt or alkaline salt stress in maize (*Zea mays* L.) seedlings. *BMC Plant Biol.* 17:41. doi: 10.1186/s12870-017-0994-6
- Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., et al. (2018). Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. *Mol. Plant* 11, 789–805. doi: 10.1016/j.molp.2018.03.018
- Gupta, B., Saha, J., Sengupta, A., and Gupta, K. (2013). Plant abiotic stress: 'omics' approach. *J. Plant Biochem. Physiol.* 1:3. doi: 10.4172/2329-9029.1000e108
- Gupta, P., and De, B. (2017). Metabolomics analysis of rice responses to salinity stress revealed elevation of serotonin, and gentisic acid levels in leaves of tolerant varieties. *Plant Signal. Behav.* 12:e1335845. doi: 10.1080/15592324
- Gupta, P. K., Mir, R. R., Mohan, A., and Kumar, J. (2008). Wheat genomics: present status and future prospects. *Int. J. Plant Genomics* 36:896451. doi: 10.1155/2008/896451
- He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H., and Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant Sci.* 5:484. doi: 10.3389/fpls.2014.00484
- Henikoff, S., Till, B. J., and Comai, L. (2004). TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* 135, 630–636. doi: 10.1104/pp.104.041061
- Hieter, P., and Boguski, M. (1997). Functional genomics: it's all how you read it. *Science (New York, NY)* 278, 601–602. doi: 10.1126/science.278.5338.601
- Hirai, M.Y., Yano, M., Goodenowe, D. B., Kanaya, S., Kimura, T., Awazuahara, M., et al. (2004). Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10205–10210. doi: 10.1073/pnas.0403218101
- Hirsch, C. N., Foerster, J. M., Johnson, J. M., Sekhon, R. S., Muttoni, G., Vaillancourt, B., et al. (2014). Insights into the maize pan-genome and pan-transcriptome. *Plant Cell* 26, 121–135. doi: 10.1105/tpc.113.119982
- Hittalmani, S., Mahesh, H. B., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y. R., et al. (2014). Genome and Transcriptome sequence of Finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18:465. doi: 10.1186/s12864-017-3850-z
- Houle, D., Govindaraju, D. R., and Omholt, S. (2010). Phenomics: the next challenge. *Nat. Rev. Genet.* 11, 855–866. doi: 10.1038/nrg2897
- Huang, D., Qiu, Y., Zhang, Y., Huang, F., Meng, J., Wei, S., et al. (2013). Fine mapping and characterization of BPH27, a brown planthopper resistance gene from wild rice (*Oryza rufipogon* Griff.). *Theor. Appl. Genet.* 126, 219–229. doi: 10.1007/s00122-012-1975-7
- Huang, L., Zhang, F., Zhang, F., Wang, W., Zhou, Y., Fu, B., et al. (2014). Comparative transcriptome sequencing of tolerant rice introgression line and its parents in response to drought stress. *BMC Genomics* 15:1026. doi: 10.1186/1471-2164-15-1026
- Huang, X.-Y., Chao, D.-Y., Gao, J.-P., Zhu, M.-Z., Shi, M., and Lin, H.-X. (2009). A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control. *Genes Dev.* 23, 1805–1817. doi: 10.1101/gad.1812409
- Huang, X. Q., Coster, H., Ganai, M. W., and Roder, M. S. (2003). Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 106, 1379–1389. doi: 10.1007/s00122-002-1179-7
- Huang, X. Y., and Salt, D. E. (2016). Plant ionomics: from elemental profiling to environmental adaptation. *Mol. Plant* 9, 787–797. doi: 10.1016/j.molp.2016.05.003
- Humlík, J. F., Lázár, D., Husíčková, A., and Spíchal, L. (2015). Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses – a review. *Plant Methods* 11:29. doi: 10.1186/s13007-015-0072-8
- Hurgobin, B., Golicz, A. A., Bayer, P. E., Chan, C. K. K., Tirnaz, S., Dolatabadian, A., et al. (2018). Parkin homeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. *Plant Biotechnol. J.* 16, 1265–1274. doi: 10.1111/pbi.12867
- Jaccoud, D., Peng, K., Feinstein, D., and Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* 29:25. doi: 10.1093/nar/29.4.e25
- Jain, S. M., and Suprasanna, P. (2011). Induced mutations for enhancing nutrition and food production. *Gene Conserve* 40, 201–215.
- Jangpromma, N., Songsri, P., Thammasirak, S., and Jaisil, P. (2010). Rapid assessment of chlorophyll content in sugarcane using a spad chlorophyll meter across different water stress conditions. *Asian J. Plant Sci.* 9, 368–374. doi: 10.3923/ajps.2010.368.374
- Jiang, W., Zhou, H., Bi, H., Fromm, M., and Yang, B. (2013). Demonstration of CRISPR/Cas9/sgrRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.* 41:e188. doi: 10.1093/nar/gkt780
- Jin, Y., Yang, H., Wei, Z., Ma, H., and Ge, X. (2013). Rice male development under drought stress: phenotypic changes and stage-dependent transcriptomic reprogramming. *Mol. Plant* 6, 1630–1645. doi: 10.1093/mp/sst067
- Jogaiah, S., Govind, S. R., and Tran, L.-S. P. (2013). Systems biology-based approaches toward understanding drought tolerance in food crops. *Crit. Rev. Biotechnol.* 33, 23–39. doi: 10.3109/07388551.2012.659174
- Johnson, S. M., Lim, F. L., Finkler, A., Fromm, H., Slabas, A. R., and Knight, M. R. (2014). Transcriptomic analysis of *Sorghum bicolor* responding to combined heat and drought stress. *BMC Genomics* 15:456. doi: 10.1186/1471-2164-15-456
- Kakumanu, A., Ambavaram, M. R., Klumas, C., Krishnan, A., Batlang, U., Myers, E., et al. (2012). Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-seq. *Plant Physiol.* 160, 846–867. doi: 10.1104/pp.112.200444
- Karkee, M., Steward, B. L., Tang, L., and Aziz, S. A. (2009). Quantifying sub-pixel signature of paddy rice field using an artificial neural network. *Comput. Electron. Agric.* 65, 65–76. doi: 10.1016/j.compag.2008.07.009
- Kawahara, Y., Oono, Y., Kanamori, H., Matsumoto, T., Itoh, T., and Minami, E. (2012). Simultaneous RNA-Seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PLoS ONE* 7:e49423. doi: 10.1371/journal.pone.0049423
- Ke, R., Marco, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., et al. (2013). *In situ* sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* 10, 857–860. doi: 10.1038/nmeth.2563
- Kearsey, M. J. (1998). The principles of QTL analysis (a minimal mathematics approach). *J. Exp. Bot.* 49, 1619–1623.
- Khan, F., Chai, H. H., Ajmera, I., Hodgman, C., Mayes, S., and Lu, C. (2017). A transcriptomic comparison of two bambara groundnut landraces under dehydration stress. *Genes* 8, 1–19. doi: 10.3390/genes8040121

- Kikuchi, J., and Hirayama, T. (2007). "Practical aspects of uniform stable isotope labeling of higher plants for heteronuclear NMR-based metabolomics," in *Metabolomics, Methods in Molecular Biology*, eds W. Weckwerth (New York, NY: Humana Press), 273–286. doi: 10.1007/978-1-59745-244-1_15
- Kim, H. K., Choi, Y. H., and Verpoorte, R. (2011). NMR-based plant metabolomics: where do we stand, where do we go? *Trends Biotechnol.* 29, 267–275. doi: 10.1016/j.tibtech.2011.02.001
- Knoll, J. E., Ramos, M. L., Zeng, Y., Holbrook, C. C., Chow, M., Chen, S., et al. (2011). TILLING for allergen reduction and improvement of quality traits in peanut (*Arachis hypogaea* L.). *BMC Plant Biol.* 11:81. doi: 10.1186/1471-2229-11-81
- Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., et al. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5:e1000551. doi: 10.1371/journal.pgen.1000551
- Kumar, A., Pathak, R. K., Gupta, S. M., Gaur, V. S., and Pandey, D. (2015a). Systems biology for smart crops and agricultural innovation: filling the gaps between genotype and phenotype for complex traits linked with robust agricultural productivity and sustainability. *OMICS* 19, 581–601. doi: 10.1089/omi.2015.0106
- Kumar, A. P. K., Boualem, A., Bhattacharya, A., Parikh, S., Desai, N., Zambelli, A., et al. (2013). SMART – sunflower mutant population and reverse genetic tools for crop improvement. *BMC Plant Biol.* 13:38. doi: 10.1186/1471-2229-13-38
- Kumar, J., Pratap, A., and Kumar, S. (2015b). "Plant phenomics: an overview," in *Phenomics in Crop Plants: Trends, Options and Limitations*, eds J. Kumar, A. Pratap, and S. Kumar, (New Delhi: Springer), 1–10. doi: 10.1007/978-81-322-2226-2_1
- Kumari, A., Das, P., Parida, A. K., and Agarwal, P. K. (2015). Proteomics, metabolomics, and ionomics perspectives of salinity tolerance in halophyte. *Front. Plant Sci.* 6:537. doi: 10.3389/fpls.2015.00537
- Kuo, T. C., Tian, T. F., and Tseng, Y. J. (2013). 3Omics: a web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data. *BMC Syst. Biol.* 7:64. doi: 10.1186/1752-0509-7-64
- Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., and Szarejko, I., et al. (2011). TILLING: a shortcut in functional genomics. *J. Appl. Genet.* 52, 371–390. doi: 10.1007/s13353-011-0061-1
- Laloum, T., Martín, G., and Duque, P. (2018). Alternative splicing control of abiotic stress responses. *Trends Plant Sci.* 23, 140–150. doi: 10.1016/j.tplants.2017.09.019
- Larrainzar, E., Wienkoop, S., Scherling, C., Kempa, S., Ladrera, R., Arrese-Igor, C., et al. (2009). Carbon metabolism and bacteroid functioning are involved in the regulation of nitrogen fixation in *Medicago truncatula* under drought and recovery. *Mol. Plant Microbe Interact.* 22, 1565–1576. doi: 10.1094/MPMI-22-12-1565
- Larrainzar, E., Wienkoop, S., Weckwerth, W., Ladrera, R., Arrese-Igor, C., and González, E. M. (2010). *Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress. *Plant Physiol.* 144, 1495–1507. doi: 10.1104/pp.107.101618
- Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., et al. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Sci. Adv.* 1:e1400218. doi: 10.1126/sciadv.1400218
- Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Ostergaard, L., Patron, N., et al. (2015). Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. *Genome Biol.* 16:258. doi: 10.1186/s13059-015-0826-7
- Le, D. T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Le, H., et al. (2012). Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by genome-wide transcriptome analysis. *PLoS ONE* 7:e49522. doi: 10.1371/journal.pone.0049522
- Le, G. H., Fontaine, J. X., Molinie, R., Pelloux, J., Mesnard, F., Gillet, F., et al. (2017). NMR-based metabolomics to study the cold-acclimation strategy of two *Miscanthus* genotypes. *Phytochem. Anal.* 28, 58–67. doi: 10.1002/pca.2649
- Li, G. Y., Zhou, J. X., Ma, W. K., Jiang, L. G., Jin, Z. H., Zhang, Y., et al. (2014). *De novo* assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nat. Biotechnol.* 32, 1045–1052. doi: 10.1038/nbt.2979
- Li, Y., Zou, W., Lin, S., Onofua, D., Yang, Z., Chen, H., et al. (2017). Transcriptome profiling and digital gene expression analysis of sweet potato for the identification of putative genes involved in the defense response against *Fusarium oxysporum* f. sp. *batatas*. *PLoS ONE* 12:e0187838. doi: 10.1371/journal.pone.0187838
- Li, Y. F., Wang, Y. I., Tang, Y., Kakani, V. G., and Mahalingam, R. (2013). Transcriptome analysis of heat stress response in switchgrass (*Panicum virgatum* L.). *BMC Plant Biol.* 13:153. doi: 10.1186/1471-2229-13-153
- Li, Z., Liu, Z. B., Xing, A., Moon, B. P., Koellhoffer, J. P., Huang, L., et al. (2015). Cas9-guide RNA directed genome editing in soybean. *Plant Physiol.* 169, 960–970. doi: 10.1104/pp.15.00783
- Lin, K., Zhang, N., Severing, E. I., Nijveen, H., Cheng, F., Visser, R. G., et al. (2014). Beyond genomic variation-comparison and functional annotation of three *Brassica rapa* genomes: a turnip, a rapid cycling and a Chinese cabbage. *BMC Genomics* 15:250. doi: 10.1186/1471-2164-15-250
- Lindon, J. C., and Nicholson, J. K. (2008). Analytical technologies for metabolomics and metabolomics, and multi-omic information recovery. *Trends Anal. Chem.* 27, 194–204. doi: 10.1016/j.trac.2007.08.009
- Liu, B., Zhang, N., Zhao, S., Chang, J., Wang, Z., Zhang, G., et al. (2015). Proteomic changes during tuber dormancy release process revealed by iTRAQ quantitative proteomics in potato. *Plant Physiol. Biochem.* 86, 181–190. doi: 10.1016/j.plaphy.2014.12.003
- Liu, D., Ford, K. L., Roessner, U., Natera, S., Cassin, A. M., Patterson, J. H., et al. (2013). Rice suspension cultured cells are evaluated as a model system to study salt responsive networks in plants using a combined proteomic and metabolomic profiling approach. *Proteomics* 13, 2046–2206. doi: 10.1002/pmic.201200425
- Lu, F., Romay, M. C., Glaubitz, J. C., Bradbury, P. J., Elshire, R. J., Wang, T., et al. (2015). High-resolution genetic mapping of maize pan-genome sequence anchors. *Nat. Commun.* 6:6914. doi: 10.1038/ncomms7914
- Lu, X., Wang, X., Chen, X., Shu, N., Wang, J., Wang, D., et al. (2017). Single-base resolution methylomes of upland cotton (*Gossypium hirsutum* L.) reveal epigenome modifications in response to drought stress. *BMC Genomics* 18:297. doi: 10.1186/s12864-017-3681-y
- Lueong, S. S., Hoheisel, J. D., and Alhamdani, M. S. S. (2014). Protein microarrays as tools for functional proteomics: achievements, promises and challenges. *J. Proteomics Bioinform.* 7:1–10. doi: 10.4172/jpb.S7-004
- Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* 24, 31–38. doi: 10.1016/j.pbi.2015.01.006
- Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4:2320. doi: 10.1038/ncomms3320
- Magalhaes, J. V., Liu, J., Guimaraes, C. T., Lana, U. G. P., Alves, V. M. C., Wang, Y. H., et al. (2007). A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat. Genet.* 39, 1156–1161. doi: 10.1038/ng2074
- Mahlein, A. K. (2016). Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. *Plant Dis.* 100, 241–251. doi: 10.1094/PDIS-03-15-0340-FE
- Maksup, S., Roytrakul, S., and Supaibulwatana, K. (2014). Physiological and comparative proteomic analyses of Thai jasmine rice and two check cultivars in response to drought stress. *J. Plant Interact.* 9, 43–55. doi: 10.1080/17429145.2012.752042
- Mangin, B., Casadebaig, P., Cadic, E., Blanchet, N., Boniface, M. C., Carrère, S., et al. (2017). Genetic control of plasticity of oil yield for combined abiotic stresses using a joint approach of crop modelling and genome-wide association. *Plant Cell Environ.* 40, 2276–2291. doi: 10.1111/pce.12961
- Margarita, P., Abbà, and, S., and Palmano, S. (2013). Novel aspects of grapevine response to phytoplasma infection investigated by a proteomic and phosphoproteomic approach with data integration into functional networks. *BMC Genomics* 14:38. doi: 10.1186/1471-2164-14-38
- Mascher, M., Gerlach, N., Gahrtz, M., Bucher, M., Scholz, U., and Dresselhaus, T. (2014). Sequence and ionomic analysis of divergent strains of maize inbred line B73 with an altered growth phenotype. *PLoS ONE* 9:e96782. doi: 10.1371/journal.pone.0096782
- Massa, A. N., Childs, K. L., and Buell, C. R. (2013). Abiotic and biotic stress responses in *Solanum tuberosum* group phureja DM1-3516R44

- as measured through whole transcriptome sequencing. *Plant Gen.* 6:15. doi: 10.3835/plantgenome2013.05.0014
- Matsuda, F., Nakabayashi, R., Yang, Z., Okazaki, Y., Yonemaru, J., Ebana, K., et al. (2015). Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. *Plant J.* 81, 13–23. doi: 10.1111/tjp.12681
- Matsui, A., Ishida, J., Morosawa, T., Mochizuki, Y., Kaminuma, E., Endo, T. A., et al. (2008). *Arabidopsis* transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol.* 49, 1135–1149. doi: 10.1093/pcp/pcn101
- Mba, C. (2013). Induced mutations unleash the potentials of plant genetic resources for food and agriculture. *Agronomy* 3, 200–231. doi: 10.3390/agronomy3010200
- McCallum, C. M., Comai, L., Greene, E. A., and Henikoff, S. (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123, 439–442. doi: 10.1104/pp.123.2.439
- McGrail, R. K., Van Sanford, D. A., and McNear, D. H. Jr. (2020). Trait-based root phenotyping as a necessary tool for crop selection and improvement. *Agronomy* 10:1328. doi: 10.3390/agronomy10091328
- McLuckey, S. A., and Stephenson, J. L. Jr. (1998). Ion/ion chemistry of high-mass multiply charged ions. *Mass Spectrom. Rev.* 17, 369–407. doi: 10.1002/(SICI)1098-2787(1998)17:6<369::AID-MAS1>3.0.CO;2-J
- Millet, E. J., Welcker, C., Kruijer, W., Negro, S., Coupel-Ledru, A., Nicolas, S. D., et al. (2016). Genome-wide analysis of yield in Europe: allelic effects vary with drought and heat scenarios. *Plant Physiol.* 172, 749–764. doi: 10.1104/pp.16.00621
- Minoia, S., Petrozza, A., D'Onofrio, O., Piron, F., Mosca, G., Sozio, G., et al. (2010). A new mutant genetic resource for tomato crop improvement by TILLING technology. *BMC Res. Not.* 3:69. doi: 10.1186/1756-0500-3-69
- Mittova, V., Guy, M., Tal, M., and Volokita, M. (2004). Salinity up-regulates the antioxidative system in root mitochondria and peroxisomes of the wild salt-tolerant tomato species *lycopersicon*. *J. Exp. Bot.* 55, 1105–1113. doi: 10.1093/jxb/erh113
- Moco, S., Bino, R. J., De Vos, R. C. H., and Vervoort, J. (2007). Metabolomics technologies and metabolite identification. *Trends Analyt. Chem.* 26, 855–866. doi: 10.1016/j.trac.2007.08.003
- Mohammadi, P. P., Moieni, A., and Komatsu, S. (2012). Comparative proteome analysis of drought-sensitive and drought-tolerant rapeseed roots and their hybrid F1 line under drought stress. *Amino Acids* 43, 2137–2152. doi: 10.1007/s00726-012-1299-6
- Montenegro, J. D., Golicz, A. A., Bayer, P. E., Hurgobin, B., Lee, H., Chan, C. K. K., et al. (2017). The pangenome of hexaploid bread wheat. *Plant J.* 90, 1007–1013. doi: 10.1111/tjp.13515
- Mosa, K. A., Ismail, A., and Helmy, M. (2017). “Omics and system biology approaches in plant stress research,” in *Plant Stress Tolerance: an Integrated Omics Approach*, eds K. A. Mosa, A. Ismail, and M. Helmy (Cham: Springer), 21–34. doi: 10.1007/978-3-319-59379-1_2
- Muthamilarasan, M., and Prasad, M. (2017). “Genetic determinants of drought stress tolerance in *Setaria*,” in *Genetics and Genomics of Setaria*. *Plant Genetics and Genomics: Crops and Models*, Vol. 19, eds A. Doust and X. Diao (Cham: Springer) 267–289. doi: 10.1007/978-3-319-45105-3_16
- Muthamilarasan, M., Singh, N. K., and Prasad, M. (2019). Multi-omics approaches for strategic improvement of stress tolerance in underutilized crop species: a climate change perspective. *Adv. Genet.* 103, 1–38. doi: 10.1016/bs.adgen.2019.01.001
- Muthurajan, R., Shobbar, Z. S., Jagadish, S., Bruskiewich, R., Ismail, A., Leung, H., et al. (2011). Physiological and proteomic responses of rice peduncles to drought stress. *Mol. Biotechnol.* 48, 173–182. doi: 10.1007/s12033-010-9358-2
- Nakagami, H., Sugiyama, N., Ishihama, Y., and Shirasu, K. (2012). Shotguns in the front line: phosphoproteomics in plants. *Plant Cell Physiol.* 53, 118–124. doi: 10.1093/pcp/pcr148
- Nataraja, K. N., Madhura, B. G., and Parvathi, S. M. (2017). “Omics: modern tools for precise understanding of drought adaptation in plants,” in *Plant OMICS and Crop Breeding*, eds S. M. Zargar and V. Rai (Palm Bay, FL: Apple Academic Press), 289–320.
- Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., and Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Sci. Rep.* 7:482. doi: 10.1038/s41598-017-00578-x
- Neveu, P., Tireau, A., Hilgert, N., Nègre, V., Mineau-Cesari, J., Brichet, N., et al. (2019). Dealing with multi-source and multi-scale information in plant phenomics: the ontology-driven phenotyping hybrid information system. *New Phytol.* 221, 588–601. doi: 10.1111/nph.15385
- Norton, G. J., Deacon, C. M., Xiong, L. Z., Huang, S. Y., Meharg, A. A., and Price, A. H. (2010). Genetic mapping of the rice ionome in leaves and grain: identification of QTLs for 17 elements including arsenic, cadmium, iron and selenium. *Plant Soil* 329, 139–153. doi: 10.1007/s11104-009-0141-8
- Nouri, M. Z., and Komatsu, S. (2010). Comparative analysis of soybean plasma membrane proteins under osmotic stress using gel-based and LC MS/MS-based proteomics approaches. *Proteomics* 10, 1930–1945. doi: 10.1002/pmic.200900632
- Novik, K. L., Nimmrich, I., Genc, B., Maier, S., Piepenbrock, C., Olek, A., et al. (2002). Epigenomics: genome-wide study of methylation phenomena. *Curr. Issues Mol. Biol.* 4, 111–128.
- Offermann, S., Danker, T., Dreymuller, D., Kalamajka, R., Topsch, S., Weyand, K., et al. (2006). Illumination is necessary and sufficient to induce histone acetylation independent of transcriptional activity at the C4-specific phosphoenolpyruvate carboxylase promoter in maize. *Plant Physiol.* 141, 1078–1088. doi: 10.1104/pp.106.080457
- Okay, S., Derelli, E., and Unver, T. (2014). Transcriptome-wide identification of bread wheat WRKY transcription factors in response to drought stress. *Mol. Genet. Genomics* 289, 765–781. doi: 10.1007/s00438-014-0849-x
- Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S. E., Kok, S. Y., Sarpan, N., et al. (2015). Loss of karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature* 525, 533–537. doi: 10.1038/nature15365
- Paine, J. A., Shipton, C. A., Chaggar, S., Howells, R. M., Kennedy, M. J., Vernon, G., et al. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat. Biotechnol.* 23, 482–487. doi: 10.1038/nbt1082
- Pandey, G., Yadav, C. B., Sahu, P. P., Muthamilarasan, M., and Prasad, M. (2017). Salinity induced differential methylation patterns in contrasting cultivars of foxtail millet (*Setaria italica* L.). *Plant Cell Rep.* 36, 759–772. doi: 10.1007/s00299-016-2093-9
- Pandey, M. K., Roorkiwal, M., Singh, V. K., Ramalingam, A., Kudapa, H., Thudi, M., et al. (2016). Emerging genomic tools for legume breeding: current status and future perspectives. *Front. Plant Sci.* 7:455. doi: 10.3389/fpls.2016.00455
- Parent, S. E., Parent, L. E., Egozcue, J. J., Rozane, D. E., Hernandez, A., Lapointe, L., et al. (2013). The plant ionome revisited by the nutrient balance concept. *Front. Plant Sci.* 4:39. doi: 10.3389/fpls.2013.00039
- Parker, D., Beckmann, M., Zubair, H., Enot, D. P., Caracul-Rios, Z., Overy, D. P., et al. (2009). Metabolomic analysis reveals a common pattern of metabolic re-programming during invasion of three host plant species by *Magnaporthe grisea*. *Plant J.* 59, 723–737. doi: 10.1111/j.1365-3113X.2009.03912.x
- Pathan, M. S., and Sleper, D. A. (2008). “Advances in soybean breeding,” in *Genetics and Genomics of Soybean*, Vol. 2, ed G. Stacey (New York, NY: Springer), 113–133. doi: 10.1007/978-0-387-72299-3_8
- Paupiere, M. J., Muller, F., Li, H. J., Rieu, I., Tikunov, Y. M., Visser, R. G. F., et al. (2017). Untargeted metabolomic analysis of tomato pollen development and heat stress response. *Plant Reprod.* 30, 81–94. doi: 10.1007/s00497-017-0301-6
- Penna, S., and Jain, S. M. (2017). Mutant resources and mutagenomics in crop plants. *Emirates J. Food Agric.* 29, 651–657. doi: 10.9755/ejfa.2017.v29.i9.86
- Pieruschka, R., Klimov, D., Kolber, Z., and Berry, J. A. (2010). Monitoring of cold and light stress impact on photosynthesis by using the laser induced fluorescence transient (LIFT) approach. *Funct. Plant Biol.* 37, 395–402. doi: 10.1071/FP09266
- Pinson, S. R. M., Tarpley, L., Yan, W., Yeater, K., Lahner, B., Yakubova, E., et al. (2015). Worldwide genetic diversity for mineral element concentrations in rice grain. *Crop Sci.* 55, 294–311. doi: 10.2135/cropsci2013.10.0656
- Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Wishart, D., et al. (2019). Systems biology and multi-omics integration: viewpoints from the metabolomics research community. *Metabolites* 9:76. doi: 10.3390/metabo9040076
- Qi, X., Xie, S., Liu, Y., Yi, F., and Yu, J. (2013). Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. *Plant Mol. Biol.* 83, 459–473. doi: 10.1007/s11103-013-0104-6
- Rabello, A. R., Guimarães, C. M., Rangel, P. H., da Silva, F. R., Seixas, D., de Souza, E., et al. (2008). Identification of drought-responsive genes in roots of upland rice (*Oryza sativa* L.). *BMC Genomics* 9:485. doi: 10.1186/1471-2164-9-485

- Rabouam, C., Comes, A. M., Bretagnolle, V., Humbert, J. F., Periquet, G., and Bigot, Y. (1999). Features of DNA fragments obtained by random amplified polymorphic DNA (RAPD) assays. *Mol. Ecol.* 8, 493–503. doi: 10.1046/j.1365-294X.1999.00605.x
- Ram, T., Majumder, N. D., Krishnaveni, D., and Ansari, M. M. (2007). Rice variety Dhanrasi, an example of improving yield potential and disease resistance by introgressing gene(s) from wild species (*Oryza rufipogon*). *Curr. Sci.* 92, 987–992.
- Ramalingam, A., Kudapa, H., Pazhamala, L. T., Weckwerth, W., and Varshney, R. K. (2015). Proteomics and metabolomics: two emerging areas for legume improvement. *Front. Plant Sci.* 6:1116. doi: 10.3389/fpls.2015.01116
- Rascher, U., and Pieruschka, R. (2008). Spatio-temporal variations of photosynthesis: the potential of optical remote sensing to better understand and scale light use efficiency and stresses of plant ecosystems. *Precis. Agric.* 9, 355–366. doi: 10.1007/s11119-008-9074-0
- Raza, A., Tabassum, J., Kudapa, H., and Varshney, R. K. (2021). Can omics deliver temperature resilient ready-to-grow crops? *Crit. Rev. Biotechnol.* 2021, 1–24. doi: 10.1080/07388551.2021.1898332
- Reynolds, D., Ball, J., Bauer, A., Davey, R., Griffiths, S., and Zhou, J. (2019). CropSight: a scalable and open-source information management system for distributed plant phenotyping and IoT-based crop management. *Gigascience* 8, 1–35. doi: 10.1093/gigascience/giz009
- Rinaldo, A. R., and Ayliffe, M. (2015). Gene targeting and editing in crop plants: a new era of precision opportunities. *Mol. Breed.* 35:40. doi: 10.1007/s11032-015-0210-z
- Ritchie, M. D., Holzinger, E. R., Li, R. W., Pendergrass, S. A., and Kim, D. (2015). Methods of integrating data to uncover genotype-phenotype interactions. *Nat. Rev. Genet.* 16, 85–97. doi: 10.1038/nrg3868
- Roitsch, T., Cabrera-Bosquet, L., Fournier, A., Ghamkhar, K., Jiménez-Berni, J., Pinto, F., et al. (2019). Review: new sensors and data-driven approaches—a path to next generation phenomics. *Plant Sci.* 282, 2–10. doi: 10.1016/j.plantsci.2019.01.011
- Ruiz-Garcia, L., Lunadei, L., Barreiro, P., and Robla, J. I. (2009). A review of wireless sensor technologies and applications in agriculture and food industry: state of the art and current trends. *Sensors* 9, 4728–4750. doi: 10.3390/s90604728
- Saand, M. A., Xu, Y. P., Li, W., Wang, J. P., and Cai, X. Z. (2015). Cyclic nucleotide gated channel gene family in tomato: genome-wide identification and functional analyses in disease resistance. *Front. Plant Sci.* 6:303. doi: 10.3389/fpls.2015.00303
- Saito, K., and Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. *Annu. Rev. Plant Biol.* 61, 463–448. doi: 10.1146/annurev.arplant.043008.092035
- Salekdeh, G. H., Siopongco, J., Wade, L. J., Ghareyazie, B., and Bennett, J. (2002). Proteomic analysis of rice leaves during drought stress and recovery. *Proteomics* 2, 1131–1145. doi: 10.1002/1615-9861(200209)2:9<1131::AID-PROT1131>3.0.CO;2-1
- Sali, A., Glaeser, R., Earnest, T., and Baumeister, W. (2003). From words to literature in structural proteomics. *Nature* 422, 216–225. doi: 10.1038/nature01513
- Salt, D. E., Baxter, I., and Lahner, B. (2008). Ionomics and the study of the plant ionome. *Annu. Rev. Plant Biol.* 59, 709–733. doi: 10.1146/annurev.arplant.59.032607.092942
- Sana, T. R., Fischer, S., Wohlgemuth, G., Katrekar, A., Jung, K. H., Ronald, P. C., et al. (2010). Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae*. *Metabolomics* 6, 451–465. doi: 10.1007/s11306-010-0218-7
- Sanchez-Rodríguez, E., Md, M., Rubio-Wilhelmi, L. M., Blasco, B., Rios, J. J., et al. (2010). Study of the ionome and uptake fluxes in cherry tomato plants under moderate water stress conditions. *Plant Soil* 335, 339–347. doi: 10.1007/s11104-010-0422-2
- Saranga, Y., Jiang, C. X., Wright, R. J., Yakir, D., and Paterson, A. H. (2004). Genetic dissection of cotton physiological responses to arid conditions and their inter-relationships with productivity. *Plant Cell Environ.* 27, 263–277. doi: 10.1111/j.1365-3040.2003.01134.x
- Satishmurti, K., Senthil, N., Vellaikumar, S., Ranjani, R. V., and Raveendran, M. (2013). Plant Ionomics: a platform for identifying novel gene regulating plant mineral nutrition. *Am. J. Plant Sci.* 4, 1309–1315. doi: 10.4236/ajps.2013.47162
- Sauter, H., Lauer, M., and Fritsch, H. (1991). “Metabolic profiling of plants – a new diagnostic technique”, in *Synthesis and Chemistry of Agrochemicals II*, eds D. R. Baker, J. G. Fenyes, and W. K. Moberg (Washington, DC: American Chemical Society), 288–299. doi: 10.1021/bk-1991-0443.ch024
- Schatz, M. C., Maron, L. G., Stein, J. C., Wences, A. H., Gurtowski, J., Biggers, E., et al. (2014). Whole genome *de novo* assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of aus and indica. *Genome Biol.* 15:506. doi: 10.1186/PREACCEPT-2784872521277375
- Schmitz, R. J., He, Y., Valdés-López, O., Khan, S. M., Joshi, T., Urlich, M. A., et al. (2013). Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res.* 23, 1663–1674. doi: 10.1101/gr.152538.112
- Schnurbusch, T., Hayes, J., and Sutton, T. (2010). Boron toxicity tolerance in wheat and barley: Australian perspectives. *Breed. Sci.* 60, 297–304. doi: 10.1270/jsbbs.60.297
- Semel, Y., Schauer, N., Roessner, U., Zamir, D., and Fernie, A. R. (2007). Metabolite analysis for the comparison of irrigated and non-irrigated field grown tomato of varying genotype. *Metabolomics* 3, 289–295. doi: 10.1007/s11306-007-0055-5
- Shah, T., Xu, J., Zou, X., Cheng, Y., Nasir, M., and Zhang, X. (2018). Omics approaches for engineering wheat production under abiotic stresses. *Int. J. Mol. Sci.* 19, 2390–2405. doi: 10.3390/ijms19082390
- Shalata, A., Mittova, V., Volokita, M., Guy, M., and Tal, M. (2001). Response of the cultivated tomato and its wild salt-tolerant relative *Lycopersicon pennellii* to salt-dependent oxidative stress: the root antioxidative system. *Physiol. Plant.* 112, 487–494. doi: 10.1034/j.1399-3054.2001.1120405.x
- Shikha, M., Kanika, A., Rao, A. R., Mallikarjuna, M. G., Gupta, H. S., and Nepolean, T. (2017). Genomic selection for drought tolerance using genome-wide SNPs in maize. *Front. Plant Sci.* 8:550. doi: 10.3389/fpls.2017.00550
- Silvente, S., Sobolev, A. P., and Lara, M. (2012). Metabolite adjustments in drought tolerant and sensitive soybean genotypes in response to water stress. *PLoS ONE* 7:e38554. doi: 10.1371/journal.pone.0038554
- Singh, U. M., Sareen, P., Sengar, R. S., and Kumar, A. (2013). Plant ionomics: a newer approach to study mineral transport and its regulation. *Acta Physiol. Plant* 35, 2641–2653. doi: 10.1007/s11738-013-1316-8
- Sirault, X. R. R., James, R. A., and Furbank, R. T. (2009). A new screening method for osmotic component of salinity tolerance in cereals using infrared thermography. *Funct. Plant Biol.* 36, 970–977. doi: 10.1071/FP09182
- Spindel, J. E., Dahlberg, J., Colgan, M., Hollingsworth, J., Sievert, J., Staggenborg, S. H., et al. (2018). Association mapping by aerial drone reveals 213 genetic associations for *Sorghum bicolor* biomass traits under drought. *BMC Genomics* 19:679. doi: 10.1186/s12864-018-5055-5
- Strahl, B., and Allis, C. (2000). The language of covalent histone modifications. *Nature* 403, 41–45. doi: 10.1038/47412
- Stroud, H., Ding, B., Simon, S. A., Feng, S., Bellizzi, M., Pellegrini, M., et al. (2013). Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife* 2:e00354. doi: 10.7554/eLife.00354
- Subba, P., Kumar, R., Gayali, S., Shekhar, S., Parveen, S., Pandey, A., et al. (2013). Characterisation of the nuclear proteome of a dehydration-sensitive cultivar of chickpea and comparative proteomic analysis with a tolerant cultivar. *Proteomics* 13, 1973–1992. doi: 10.1002/pmic.201200380
- Sumner, L. W., Mendes, P., and Dixon, R. A. (2003). Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* 62, 817–836. doi: 10.1016/s0031-9422(02)00708-2
- Sun, C. X., Gao, X. X., Li, M. Q., Fu, J. Q., and Zhang, Y. L. (2016). Plastic responses in the metabolome and functional traits of maize plants to temperature variations. *Plant Biol.* 18, 249–261. doi: 10.1111/plb.12378
- Sun, X., and Weckwerth, W. (2012). COVAIN: a toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 8, 81–93. doi: 10.1007/s11306-012-0399-3
- Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., et al. (2008). MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* 279, 213–223. doi: 10.1007/s00438-007-0293-2
- Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C., and Cigan, A. M. (2015). Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* 169, 931–945. doi: 10.1104/pp.15.00793

- Swarbrick, P. J., Schulze-Lefert, P., and Scholes, J. D. (2006). The metabolic consequences of susceptibility and the activation of race specific or broad spectrum resistance pathways in barley leaves challenged with the powdery mildew fungus. *Plant Cell Environ.* 29, 1061–1076. doi: 10.1111/j.1365-3040.2005.01472.x
- Talukdar, D., and Sinjushin, A. (2015). “Cytogenomics and mutagenomics in plant functional biology and breeding,” in *PlantOmics: The Omics of Plant Science*, eds D. Barh, M. Khan, and E. Davies (New Delhi: Springer), 113–156. doi: 10.1007/978-81-322-2172-2_5
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T. (1988). Protein and polymer analyses up to m/z 100000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 2, 151–153. doi: 10.1002/rcm.1290020802
- Tao, Y., Mace, E., George-Jaeggli, B., Hunt, C., Cruickshank, A., Henzell, R., et al. (2018). Novel grain weight loci revealed in a cross between cultivated and wild sorghum. *Plant Genome* 11, 1–10. doi: 10.3835/plantgenome2017.10.0089
- Tao, Y., Mace, E. S., Tai, S., Cruickshank, A., Campbell, B. C., Zhao, X., et al. (2017). Whole-genome analysis of candidate genes associated with seed size and weight in *sorghum bicolor* reveals signatures of artificial selection and insights into parallel domestication in cereal crops. *Front. Plant Sci.* 8:1237. doi: 10.3389/fpls.2017.01237
- Tettelin, H., Massignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial ‘pan-genome’. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13950–13955. doi: 10.1073/pnas.0506758102
- Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., et al. (2004). Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* 4:12. doi: 10.1186/1471-2229-4-12
- Todaka, D., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2015). Recent advances in the dissection of drought-stress regulatory networks and strategies for development of drought-tolerant transgenic rice plants. *Front. Plant Sci.* 6:84. doi: 10.3389/fpls.2015.00084
- Tokimatsu, T., Sakurai, N., Suzuki, H., Ohta, H., Nishitani, K., Koyama, T., et al. (2005). KaPPA-view: a web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiol.* 138, 1289–1300. doi: 10.1104/pp.105.060525
- Tomlekova, N. B. (2010). Induced mutagenesis for crop improvement in Bulgaria. *Plant Mutat. Rep.* 2, 4–27.
- Toorchi, M., Yukawa, K., Nouri, M. Z., and Komatsu, S. (2009). Proteomics approach for identifying osmotic-stress-related proteins in soybean roots. *Peptides* 30, 2108–2117. doi: 10.1016/j.peptides.2009.09.006
- Twyman, R. M. (2013). *Principles of Proteomics*. Abingdon: Garland Science Press.
- Urano, K., Kurihara, Y., Seki, M., and Shinozaki, K. (2010). “Omics” analyses of regulatory networks in plant abiotic stress responses. *Curr. Opin. Plant Biol.* 13, 132–113. doi: 10.1016/j.pbi.2009.12.006
- Urbanczyk-Wochniak, E., Luedemann, A., Kopka, J., Selbig, J., Roessner-Tunalı, U., Willmitzer, L., et al. (2003). Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO Rep.* 4, 989–993. doi: 10.1038/sj.embor.embor944
- van Dijk, K., Ding, Y., Malkaram, S., Riethoven, J. J. M., Liu, R., Yang, J., et al. (2010). Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in response to dehydration stress in *Arabidopsis thaliana*. *BMC Plant Biol.* 10:238. doi: 10.1186/1471-2229-10-238
- Varshney, R. K., Thudi, M., May, G. D., and Jackson, S. A. (2010). Legume genomics and breeding. *Plant Breed. Rev.* 33, 257–304. doi: 10.1002/9780470535486.ch6
- Viant, M. R., and Sommer, U. (2013). Mass spectrometry based environmental metabolomics: a primer and review. *Metabolomics* 9, 144–158. doi: 10.1007/s11306-012-0412-x
- Vos, P., Hogers, R., Bleeker, M., Reijmans, M., Lee, T., Hornes, M., et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23, 4407–4414. doi: 10.1093/nar/23.21.4407
- Wang, M., Wang, P., Tu, L., Zhu, S., Zhang, L., Li, Z., et al. (2016). Multi-omics maps of cotton fibre reveal epigenetic basis for staged single-cell differentiation. *Nucleic Acids Res.* 44, 4067–4079. doi: 10.1093/nar/gkw238
- Wang, W. S., Mauleon, R., Hu, Z. Q., Chebotarov, D., Tai, S. S., Wu, Z. C., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557, 43–49. doi: 10.1038/s41586-018-0063-9
- Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., et al. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951. doi: 10.1038/nbt.2969
- Weckwerth, W. (2003). Metabolomics in systems biology. *Annu. Rev. Plant Biol.* 54, 669–689. doi: 10.1146/annurev.arplant.54.031902.135014
- Weckwerth, W. (2010). Metabolomics: an integral technique in systems biology. *Bioanalysis* 2, 829–836. doi: 10.4155/bio.09.192
- Weckwerth, W. (2011a). Green systems biology – from single genomes, proteomes and metabolomes to ecosystems research and biotechnology. *J. Proteome* 75, 284–305. doi: 10.1016/j.jpro.2011.07.010
- Weckwerth, W. (2011b). Unpredictability of metabolism—the key role of metabolomics science in combination with next-generation genome sequencing. *Anal. Bioanal. Chem.* 400:1967. doi: 10.1007/s00216-011-4948-9
- Weckwerth, W. (2019). Toward a unification of system-theoretical principles in biology and ecology—the stochastic lyapunov matrix equation and its inverse application. *Front. Appl. Math. Stat.* 5:29. doi: 10.3389/fams.2019.00029
- Weckwerth, W., and Fiehn, O. (2002). Can we discover novel pathways using metabolomic analysis? *Curr. Opin. Biotechnol.* 13, 156–160. doi: 10.1016/s0958-1669(02)00299-9
- Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P., and Varshney, R. K. (2020). PANOMICS meets germplasm. *Plant Biotechnol. J.* 10:1111. doi: 10.1111/pbi.13372
- Weckwerth, W., and Morgenthal, K. (2005). Metabolomics: from pattern recognition to biological interpretation. *Drug Discov Today* 10, 1551–1155. doi: 10.1016/S1359-6446(05)03609-3
- Wen, W., Li, D., Li, X., Gao, Y. Q., Li, W. Q., Li, H. H., et al. (2014). Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights. *Nat. Commun.* 5:3438. doi: 10.1038/ncomms4438
- Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acid Res.* 18, 6531–6535. doi: 10.1093/nar/18.22.6531
- Witt, S., Galicia, L., Lisec, J., Cairns, J., Tieszen, A., Araus, J. L., et al. (2012). Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress. *Mol. Plant* 5, 401–417. doi: 10.1093/mp/ssr102
- Witzel, K., Neugart, S., Ruppel, S., Schreiner, M., Wiesner, M., and Baldermann, S. (2015). Recent progress in the use of ‘omics technologies in brassicaceous vegetables. *Front. Plant Sci.* 6:244. doi: 10.3389/fpls.2015.00244
- Wohlbach, D. J., Quirino, B. F., and Sussman, M. R. (2008). Analysis of the *Arabidopsis* histidine kinase *ATHK1* reveals a connection between vegetative osmotic stress sensing and seed maturation. *Plant Cell* 20, 1101–1117. doi: 10.1105/tpc.107.055871
- Woolfson, M. (2018). The development of structural x-ray crystallography. *Phys. Scr.* 93, 1–32. doi: 10.1088/1402-4896/aa9c30
- Wu, D., Shen, Q., Cai, S., Chen, Z. H., Dai, F., and Zhang, G. (2013). Ionic responses and correlations between elements and metabolites under salt stress in wild and cultivated barley. *Plant Cell Physiol.* 54, 1976–1988. doi: 10.1093/pcp/pct134
- Wu, S., Ning, F., Zhang, Q., Wu, X., and Wang, W. (2017). Enhancing omics research of crop responses to drought under field conditions. *Front. Plant Sci.* 8:174. doi: 10.3389/fpls.2017.00174
- Xiao, J. H., Li, J. M., Grandillo, S., Ahn, S. N., Yuan, L. P., Tanksley, S. D., et al. (1998). Identification of trait-improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics* 150, 899–909.
- Xiong, L., Schumaker, K. S., and Zhu, J. K. (2002). Cell signaling during cold, drought, and salt stress. *Plant Cell* 14, 65–S183. doi: 10.1105/tpc.000596
- Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F. Y., Li, X. Y., et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nat. Biotechnol.* 30, 105–111. doi: 10.1038/nbt.2050
- Yadav, C. B., Pandey, G., Muthamilarasan, M., and Prasad, M. (2018). “Epigenetics and epigenomics of plants,” in *Plant Genetics and Molecular Biology. Advances in Biochemical Engineering/Biotechnology*, Vol. 64, eds R. Varshney, M. Pandey, and A. Chitkineni (Cham: Springer), 237–261. doi: 10.1007/10_2017_51
- Yang, F., Melo-Braga, M. N., Larsen, M. R., Jørgensen, H. J. L., and Palmisano, G. (2013). Battle through signaling between wheat and the fungal pathogen

- Septoria tritici* revealed by proteomics and phosphoproteomics. *Mol. Cell. Proteomics* 12, 2497–2508. doi: 10.1074/mcp.M113.027532
- Yang, L. N., Peng, L., Zhang, L. M., Zhang, L. L., and Yang, S. S. (2009). A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on Back propagation artificial neural network and principal components analysis. *Comput. Electron. Agric.* 68, 200–206. doi: 10.1016/j.compag.2009.06.003
- Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5:5087. doi: 10.1038/ncomms6087
- Yang, W., Guo, Z., Huang, C., Wang, K., Jiang, N., Fen, H., et al. (2015). Genome wide association study of rice (*Oryza sativa* L.) leaf traits with a high-throughput leaf scorer. *J. Exp. Bot.* 66, 5605–5615. doi: 10.1093/jxb/erv100
- Yang, Y., Saand, M. A., Abdelaal, W. B., Zhang, J., Wu, Y. i., Li, J., et al. (2020). TRAQ-based comparative proteomic analysis of two coconut varieties reveals aromatic coconut cold-sensitive in response to low temperature. *J. proteomics* 220:103766. doi: 10.1016/j.jprot.2020.103766
- Yu, J., Holland, J. B., McMullen, M. D., and Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics* 178, 539–551. doi: 10.1534/genetics.107.074245
- Zhang, C., Yang, H., and Yang, H. (2015). Evolutionary character of alternative splicing in plants. *Bioinform. Biol. Insights* 9, 47–52. doi: 10.4137/BBLI33716
- Zhang, M., Lv, D., Ge, P., Bian, Y., Chen, G., and Zhu, G. (2014a). Phosphoproteome analysis reveals new drought response and defense mechanisms of seedling leaves in bread wheat (*Triticum aestivum* L.). *J. Proteomics* 109, 290–308. doi: 10.1016/j.jprot.2014.07.010
- Zhang, M., Pinson, S. R., Tarpley, L., Huang, X. Y., Lahner, B., Yakubova, E., et al. (2014b). Mapping and validation of quantitative trait loci associated with concentrations of 16 elements in unmilled rice grain. *Theor. Appl. Genet.* 127, 137–165. doi: 10.1007/s00122-013-2207-5
- Zhang, X., Huang, C., Wu, D., Qiao, F., Li, W., Duan, L., et al. (2017). High throughput phenotyping and QTL mapping reveals the genetic architecture of maize plant growth. *Plant Physiol.* 173, 1554–1564. doi: 10.1104/pp.16.01516
- Zhang, X., Zhou, S. X., Fu, Y. C., Su, Z., Wang, X. K., and Sun, C. Q. (2006). Identification of a drought tolerant introgression line derived from Dongxiang common wild rice (*O-rufipogon* Griff.). *Plant Mol. Biol.* 62, 247–259. doi: 10.1007/s11103-006-9018-x
- Zhao, C., Zhang, Y., Du, J., Guo, X., Wen, W., Gu, S., et al. (2019). Crop phenomics: current status and perspectives. *Front. Plant Sci.* 10:714. doi: 10.3389/fpls.2019.00714
- Zhao, Q., Feng, Q., Lu, H. Y., Li, Y., Wang, A., Tian, Q. L., et al. (2018). Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nat. Genet.* 50, 278–284. doi: 10.1038/s41588-018-0041-z
- Zhong, L., Xu, Y., and Wang, J. (2009). DNA-methylation changes induced by salt stress in wheat *Triticum aestivum*. *Afr. J. Biotechnol.* 8, 6201–6207. doi: 10.5897/AJB09.1058
- Zhong, S., Fei, Z., Chen, Y. R., Zheng, Y., Huang, M., Vrebalov, J., et al. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* 31, 154–159. doi: 10.1038/nbt.2462
- Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 408–414. doi: 10.1038/nbt.3096
- Zhu, H. G., Cheng, W. H., Tian, W. G., Li, Y. J., Liu, F., Xue, F., et al. (2018). iTRAQ-based comparative proteomic analysis provides insights into somatic embryogenesis in *Gossypium hirsutum* L. *Plant Mol. Biol.* 96, 89–102. doi: 10.1007/s11103-017-0681-x
- Zhu, Y. N., Shi, D. Q., Ruan, M. B., Zhang, L. L., Meng, Z. H., Liu, J., et al. (2013). Transcriptome analysis reveals crosstalk of responsive genes to multiple abiotic stresses in cotton (*Gossypium hirsutum* L.). *PLoS ONE* 8:e80218. doi: 10.1371/journal.pone.0080218
- Ziegler, G., Terauchi, A., Becker, A., Armstrong, P., Hudson, K., and Baxter, I. (2013). Ionomics screening of field-grown soybean identifies mutants with altered seed elemental composition. *Plant Gen.* 6, 1–9. doi: 10.3835/plantgenome2012.07.0012
- Zong, W., Zhong, X., You, J., and Xiong, L. (2013). Genome-wide profiling of histone H3K4-trimethylation and gene expression in rice under drought stress. *Plant Mol. Biol.* 81, 175–188. doi: 10.1007/s11103-012-9990-2
- Zuo, W. L., Chao, Q., Zhang, N., Ye, J. R., Tan, G. Q., Li, B. L., et al. (2015). A maize wall-associated kinase confers quantitative resistance to head smut. *Nat. Genet.* 47, 151–157. doi: 10.1038/ng.3170

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Yang, Saand, Huang, Abdelaal, Zhang, Wu, Li, Sirohi and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Pearl Millet: A Climate-Resilient Nutricereal for Mitigating Hidden Hunger and Provide Nutritional Security

C. Tara Satyavathi^{1*}, Supriya Ambawat¹, Vikas Khandelwal¹ and Rakesh K. Srivastava²

¹ Indian Council of Agricultural Research - All India Coordinated Research Project on Pearl Millet, Jodhpur, India,

² Department of Molecular Breeding (Genomics Trait Discovery), International Crops Research Institute for Semi-arid Tropics, Patancheru, India

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Mehanathan Muthamilarasan,
University of Hyderabad, India
Arindam Ghatak,
University of Vienna, Austria
Ajit Singh,
University of Nottingham Malaysia
Campus, Malaysia

*Correspondence:

C. Tara Satyavathi
csatyavathi@gmail.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 28 January 2021

Accepted: 03 August 2021

Published: 13 September 2021

Citation:

Satyavathi CT, Ambawat S,
Khandelwal V and Srivastava RK
(2021) Pearl Millet: A Climate-Resilient
Nutricereal for Mitigating Hidden
Hunger and Provide Nutritional
Security. *Front. Plant Sci.* 12:659938.
doi: 10.3389/fpls.2021.659938

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is the sixth most important cereal crop after rice, wheat, maize, barley and sorghum. It is widely grown on 30 million ha in the arid and semi-arid tropical regions of Asia and Africa, accounting for almost half of the global millet production. Climate change affects crop production by directly influencing biophysical factors such as plant and animal growth along with the various areas associated with food processing and distribution. Assessment of the effects of global climate changes on agriculture can be helpful to anticipate and adapt farming to maximize the agricultural production more effectively. Pearl millet being a climate-resilient crop is important to minimize the adverse effects of climate change and has the potential to increase income and food security of farming communities in arid regions. Pearl millet has a deep root system and can survive in a wide range of ecological conditions under water scarcity. It has high photosynthetic efficiency with an excellent productivity and growth in low nutrient soil conditions and is less reliant on chemical fertilizers. These attributes have made it a crop of choice for cultivation in arid and semi-arid regions of the world; however, fewer efforts have been made to study the climate-resilient features of pearl millet in comparison to the other major cereals. Several hybrids and varieties of pearl millet were developed during the past 50 years in India by both the public and private sectors. Pearl millet is also nutritionally superior and rich in micronutrients such as iron and zinc and can mitigate malnutrition and hidden hunger. Inclusion of minimum standards for micronutrients—grain iron and zinc content in the cultivar release policy—is the first of its kind step taken in pearl millet anywhere in the world, which can lead toward enhanced food and nutritional security. The availability of high-quality whole-genome sequencing and re-sequencing information of several lines may aid genomic dissection of stress tolerance and provide a good opportunity to further exploit the nutritional and climate-resilient attributes of pearl millet. Hence, more efforts should be put into its genetic enhancement and improvement in inheritance to exploit it in a better way. Thus, pearl millet is the next-generation crop holding the potential of nutritional richness and the climate resilience and efforts must be targeted to develop nutritionally dense hybrids/varieties tolerant to drought using different omics approaches.

Keywords: pearl millet, climate-resilience, drought tolerance, abiotic stress, nutritional security

INTRODUCTION

The changing climate is leading to an increase in global average temperature affecting agricultural production worldwide. Further, it directly influences biophysical factors such as plant and animal growth along with the different areas associated with food processing and distribution. Assessment of effects of global climate changes and deployment of new tools and strategies to mitigate their effect is crucial to maximizing agricultural production to meet out food demands of the increasing population. In this context, pearl millet is most useful as it is a nutritious, climate change-ready crop with enormous potential for yielding higher economic returns in marginal conditions in comparison with other cereals even in case of climate change with harsh temperature conditions. Moreover, it has greater ceiling temperatures for grain yield and is an underutilized crop with huge nutritional potential, which needs to be utilized fully (Krishnan and Meera, 2018). It is more resilient to extreme climatic events such as drought and water scarcity and can play a vital role in ensuring food and nutritional security in changing climatic scenarios, which is mounting to frightening proportions. Globally, it is the sixth most significant cereal crop after rice (*Oryza sativa*), wheat (*Triticum aestivum*), maize (*Zea mays*), barley (*Hordeum vulgare*) and sorghum (*Sorghum bicolor*). It is a staple food of 90 million poor people and extensively grown on 30-million-ha area in the arid and semi-arid tropical regions of Asia and Africa. It is also used for feed and fodder and accounts for almost half of the global millet production (Srivastava et al., 2020a). It is mainly cultivated on marginal lands under rainfed conditions and can sustain and produce a significant amount of grain even in drought-prone areas that receive an average annual precipitation of <250 mm (Nambiar et al., 2011). It surpasses all other cereals such as wheat, maize, rice, sorghum and barley because of its unique attributes like the C₄ plant having high photosynthetic efficiency, more dry matter production capability, and survival under adverse agro-climatic conditions with lesser inputs and more economic returns (Nambiar et al., 2011). C₄ plants have more ability to fix inorganic CO₂ and more efficient in water utilization in comparison with C₃ plants due to the presence of “Kranz” anatomy in leaves. Thus, being a C₄ plant, pearl millet can account for 30% of global terrestrial carbon fixation along with other C₄ plants such as maize and sorghum (Choudhary et al., 2020). It also possesses several advantages such as early maturity, drought tolerance, the requirement of minimal inputs, and usually free from biotic and abiotic stresses. Its inherent ability to endure high temperatures up to 42°C during the reproductive phase makes it suitable for growth in extremely hot summers under irrigations in northern Gujarat and eastern Uttar Pradesh of India, thus making it a climate-resilient crop.

It also possesses the huge capability to eliminate micronutrient deficiency among developing countries (Rai et al., 2012; Anuradha et al., 2017; Singhal et al., 2018) as it supplies 30–40% of inorganic nutrients and bestows affordable staple food along with ample amounts of iron and zinc (Rao et al., 2006). It has very high nutritional values and is a good source of energy, carbohydrates, crude fibers [resistant starch (RS), soluble and

insoluble dietary fibers], soluble and insoluble fat, proteins (8–19%), ash, dietary fibers (1.2 g/100 g), antioxidants and fat (3–8%) with better fat digestibility, iron, and zinc in comparison with other major cereals (Uppal et al., 2015). It is also a rich source of vitamins such as riboflavin, niacin, and thiamine and minerals (2.3 mg/100 g) such as potassium, phosphorous, magnesium, iron, zinc, copper, and manganese (Weckwerth et al., 2020). It exhibits a better essential amino acid profile of protein in comparison with other cereals such as maize and rice. It contains lesser cross-linked prolamins leading to higher digestibility of the millet proteins. It has 74% polyunsaturated fatty acids (PUFAs) and rich in nutritionally sought-after omega-3 fatty acids such as oleic acid (25%), linoleic acid (45%), and linolenic acid (4%), which are considered best for health (Rooney, 1978; Nantanga, 2006; Dyll, 2015; Singh et al., 2018). It is a gluten-free grain that retains alkaline properties even after being cooked and is thus good for people suffering from gluten allergy. It owns a higher quantity of slowly digestible starch (SDS) and RS that account for lower glycemic index (GI) and is much preferred in recent times of transforming diets, food habits and the food industry (Satyavathi et al., 2020). It is a highly nutritious, non-acid-forming, non-glutinous food having several nutraceuticals and health beneficial properties along with high fiber content. It acts as a probiotic food for microflora present in our body and keeps us away from constipation. It is also capable of lowering cholesterol due to the presence of niacin in its grain. It contributes to an antioxidant activity with phytates, polyphenols etc. Consumption of various types of millets is considered to protect against certain types of cancer, cardiovascular diseases and various age-related diseases. Due to these useful properties, pearl millet is gaining a lot of popularity among health-conscious people all over the world. Due to its nutritional properties, pearl millet has been renamed as *nutri-cereal* (Gazette of India, No. 133 dated 13 April, 2018) and can play a vital role in overcoming malnutrition and ensure food and nutritional security.

In the present-changing climatic scenario, abiotic stresses entail a huge risk for plant growth and development leading to an over 50% decrease in the yield among the popular cereal crops (Bray et al., 2000). Almost 90% of the cultivable land is affected by various abiotic stresses globally, while only 10% of the agricultural land is free from these abiotic stresses (Dita et al., 2006). Drought and heat are the two most significant production constraints existing among the different environmental stresses. In this context, a crop species like pearl millet, which is resilient to higher temperatures and lower rainfall, can play a crucial role in fulfilling the increasing food demands of the growing population of the world. Pearl millet is mainly cultivated on marginal lands facing untimely and irregular rainfall patterns and environmental stresses due to its natural inbuilt capacity to survive in such areas and withstand abiotic stresses such as drought, salinity, heat etc. (Serba and Yadav, 2016). Genetic improvement of pearl millet through increased production was realized using hybrid technology and conventional breeding methods of selection (Yadav and Rai, 2013; Yadav et al., 2021) but later, various biotechnological and genomic approaches were used for further improvement (Varshney et al., 2017; Bollam et al., 2018; Ambawat et al., 2020). Genetic maps, next-generation

sequencing (NGS), genotyping-by-sequencing (GBS), genome-wide association studies (GWAS), synteny studies, expression profiling, fine QTL mapping, candidate gene identification and genetic engineering, gene pyramiding, bioinformatics and systems biology are some of the useful platforms, which are further being used for the genetic improvement of this nutricereal. Recently reported genome sequence information in the year 2017 can speed up gene innovation and trait mapping and can help in the understanding of several complicated gene pathways and their interactions (Varshney et al., 2017). Further, the challenge remains to characterize thousands of genes crucial for abiotic stress response and tolerance. Similarly, there is a high need to identify the lines that can use nitrogen efficiently as most of the agro-ecologies where pearl millet is grown have low N in the root zone soil strata.

In addition, integrated knowledge on genomics as well as transcriptomics, proteomics and metabolomics is also beneficial for advancements and biofortification of pearl millet (Dita et al., 2006; Lata, 2015; Anuradha et al., 2017; Ambawat et al., 2020). Hence, there is a need to focus on this very important crop and harness its suitability to adverse conditions and utilize its inbuilt capacity to ensure global food and nutritional security. Here, we have reviewed the importance of pearl millet in the present-changing climatic scenario for food and nutritional security and various advances made in the pearl millet improvement program.

STATUS OF PEARL MILLET PRODUCTION

Pearl millet is a descendent of the wild West African grass and was domesticated over 4,000 years ago in the West African Sahel, spreading later to East Africa and India (Sharma et al., 2021). Now it is being cultivated over 30 million ha worldwide, with the majority of the crop grown in Africa (>18 million ha) and Asia (>10 million ha) (Raheem et al., 2021). Around 90 million people in the Sahelian region of Africa and northwestern India consume pearl millet grain as a staple food (Srivastava et al., 2020b). Jukanti et al. (2016) have reviewed the origin and evolutionary history of pearl millet. It is the sixth major cereal crop in the world followed by maize, rice, wheat, barley and sorghum and cultivated on 30 million ha in the arid and semi-arid tropical regions of Asia and Africa accounting for around half of the global millet production with 60% of the cultivation area in Africa, followed by 35% in Asian countries. In terms of area and production, India is the largest producer of pearl millet. During 2010–2012, the average pearl millet area in India was 8.5 million ha and the average production was 9.4 million tons. It is taken up in an area of 6.93 million ha with an average production of 8.61 million tones and 1,243 kg/ha productivity (Directorate of Millets Development, 2020). The trends of area, production, and productivity over the years are shown in **Figure 1**. In Africa, the west/central Africa (WCA) region (Nigeria, Niger, Chad, Mali, and Senegal), and east/southern Africa (ESA), which includes Sudan, Ethiopia, and Tanzania, are the two main areas of pearl millet cultivation. Pearl millet is the third major crop in sub-Saharan Africa with Nigeria, Senegal, Chad, Mali, Niger and Burkina Faso as the major producing countries and has socio-economic, food/feed,

health and environmental impact on the resource-poor people of Africa. WCA is the largest pearl millet-producing region in Africa and the world, accounting for 95 % of the total area in WCA (Jukanti et al., 2016). West Africa is the largest producer led by Nigeria (41%), Niger (16%), Burkina Faso (7%), Mali (6.4%), and Senegal and Sudan (4.8%). In Africa, it is produced in 18.50 million ha by 28 countries with a yield of 11.36 million tons covering 30% different areas of the continent in diverse agro-ecologies. It is 49% of the global millet area with great significance (FAO, 2019).

It is extensively cultivated in India and is the fourth most extensively grown cereal crop after rice, wheat and maize. Rajasthan, Maharashtra, Uttar Pradesh, Gujarat, and Haryana are the major pearl millet-growing states, contributing 90% of the total production in India. Out of this, Rajasthan contributes a maximum of around 4.283 million 5 tonnes, followed by Uttar Pradesh (1.302), Haryana (1.079), Gujarat (0.961), Maharashtra (0.66), and Tamil Nadu (0.084). It is mainly cultivated in the rainy (*kharif*) season (June/July–September/October) but it is also grown in some parts of Gujarat, Rajasthan, and Uttar Pradesh during the summer season (February–May), while it is also cultivated in states of Maharashtra and Gujarat at a small scale during the post-rainy (*rabi*) season (November–February). As millets are climate-smart crops with nutritional value, they are rightly termed as *nutricereals* (Gazette of India, No. 133 dated 13 April, 2018). In addition, to include millets into the mainstream and exploit its nutritionally superior qualities and promote its cultivation, Government of India has declared Year 2018 as the “Year of Millets” and FAO Committee on Agriculture (COAG) forum has declared Year 2023 as “International Year of Millets.”

PEARL MILLET IMPROVEMENT

Indian Council of Agricultural Research started pearl millet breeding in India in the 1940s, and X1 and X2 were the two chance hybrids released in India for commercial utilization in the fifties (Yadav and Rai, 2013). Pearl millet improvement programs were implemented in several phases. During phase I, breeders mainly focused on the flowering habit, mode of pollination, germplasm evaluation and enhancement, genetics and cytogenetics of agronomically important traits, cytoplasmic male sterility (CMS) etc. Thus, initially, efforts were put forward towards the identification and use of dwarfing genes for enhancing the yield using locally adapted materials and various Open Pollinated Varieties (OPVs) were developed. As a result of this, pearl millet hybrid research has gained importance in India and the productivity was 4.5 kg/ha/year during this phase (Yadav et al., 2019). By the 1960s, hybrid development became the major objective of breeding for enhancing pearl millet production and productivity. Hybrid “HB-1” (Hybrid Bajra-1) was the first pearl millet hybrid released in 1965 (Athwal, 1965) followed by a series of hybrids between 1965 and 1988 and during phase II, an annual increase of 6.6 kg/ha was achieved in productivity.

The genetic improvement program progressed effectively initiating from the selection of local and traditional germplasm to the development of high-yielding hybrids possessing inbuilt

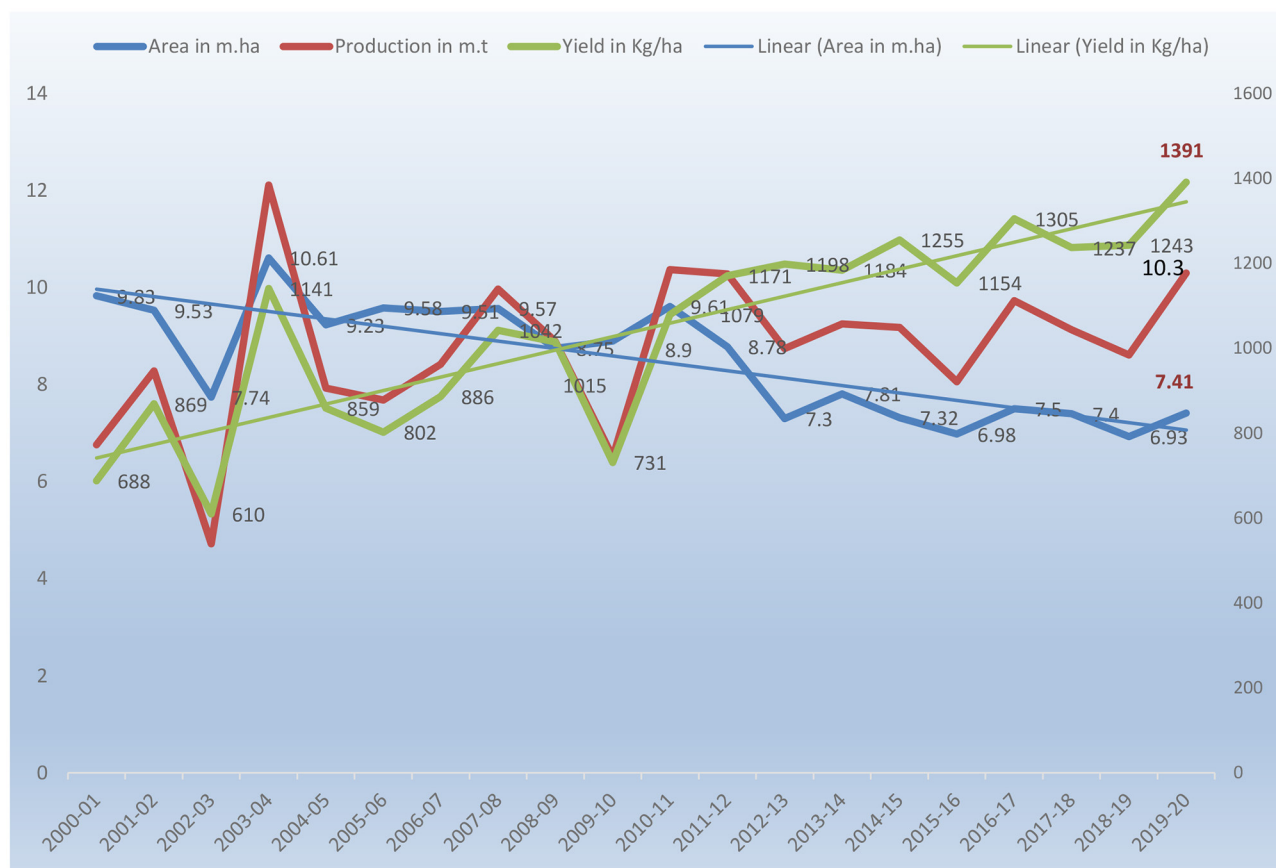


FIGURE 1 | Area, production and productivity of Pearl millet in India since 2000.

tolerance and resistance to climatic stresses such as drought and heat along with different diseases. These hybrids were grown on 70% of the total pearl millet area, resulting in a 124% enhancement in productivity since 1986–1990. ICAR-All India Coordinated Research Project on Pearl Millet has developed several precise production and protection technologies for different agro-ecological regions of different states since its beginning in 1965. Till now, a total of 180 hybrids and 62 varieties have been identified and released for growing in different agro-ecological regions of India through ICAR-All India Coordinated Research Project on Pearl millet (Satyavathi et al., 2020). Several genetically diverse CMS lines have been developed and used along with marker-assisted breeding (MAB) and marker-assisted backcrossing, resulting in an increased productivity to 19.0 kg/ha/year during phase III (Yadav et al., 2019).

During the fourth phase, great emphasis was laid on the genetic diversity of seed and pollinator parents and adaptation to niche areas, resulting in the release of a large number of cultivars and a significant increase in productivity 31.1 kg/ha/year, which was almost five times in comparison with Green Revolution Era (Govindaraj et al., 2010; Kumara et al., 2014; Yadav et al., 2019). In the next phase, biofortification of the grain for micronutrients, largely for zinc and iron and application of molecular techniques

were focused to speed up the cultivar development program (Rai et al., 2013; Kanatti et al., 2016; Kumar et al., 2016, 2018; Anuradha et al., 2017; Singhal et al., 2018; Govindaraj et al., 2019). These cultivars were widely adopted by Indian farmers resulting in enhanced crop productivity from 305 kg/ha during 1951–1955 to 998 kg/ha during 2008–2012 and 1,243 kg/ha during 2018–19 (Yadav and Rai, 2013; Satyavathi et al., 2020) (Figure 2). The productivity during the 5 years starting from 1951 to 1955 (305 kg/ha) has increased to 1,290 kg/ha for the 5 years 2016–2020. The productivity improvement is four-fold, or 400%. Compared to the previous 5 years term of 2011–2015, the productivity was 1,192 kg/ha, while it is 1,290 kg/ha for the 5 years 2016–2020. The improvement is 7.6% compared to the period of 2011–2015. This improvement is due to the combined contribution of the development of high-yielding hybrids, varieties, biofortified genotypes, improved production practices, technologies, and recommendations coupled with adoption by farmers.

Several efforts were taken for crop improvement in pearl millet using conventional as well as advanced molecular and genomic tools as listed in Tables 1, 2. Pearl millet is the first crop in which marker-assisted-selection (MAS) strategies and tools were applied to get improved varieties. Yadav et al. (2021) discussed various past strategies and future approaches

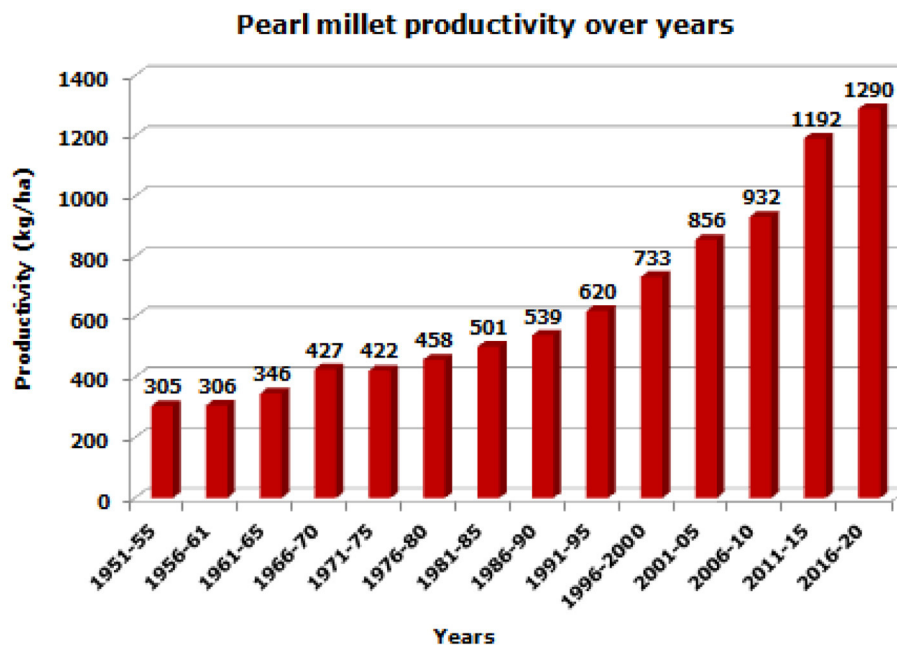


FIGURE 2 | Trends in pearl millet productivity over years (based on 5-year average).

to accelerate genetic gains to meet future demand. Further, they also emphasized the importance of genome editing, pre-breeding, precision phenotyping protocols and speed breeding approaches for pearl millet improvement and enhanced genetic gains. On the other hand, due to a lack of knowledge and non-availability of ideal dose of fertilizer among poor farmers, it becomes difficult to harvest the real yield potential. Hence, the yield improvement of pearl millet under low nitrogen input is indeed beneficial for economic and environmentally sustainable cultivation. Pujarula et al. (2021) studied the genetic variation for nitrogen-use efficiency (NUE) among a set of 380 diverse pearl millet lines. Thus, such studies of different physiological traits and their relationship with grain yield are very important for understanding the complex nature of NUE.

CLIMATE RESILIENCE IN PEARL MILLET AND GENOMIC RESOURCES

Pearl millet can survive and produce a large quantity of grain, whereas other cereals such as rice, wheat, maize, sorghum and barley may fail to provide economic benefits under adverse conditions and poor soil. It can provide multiple securities in form of food, fodder, livelihood, nutrition health and ecological benefits, whereas wheat and rice provide only food security, thus making it a crop of agricultural security. Its ability to withstand higher temperature and survival in drought-prone areas and cultivation in parts of Gujarat and eastern Uttar Pradesh of India during hot summers makes it a climate-resilient crop for overcoming the adverse effects of the changing climate (Gupta et al., 2015). National Agricultural Research

System (NARS) in India and International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) have played a significant role in developing various improved breeding and parental lines of prospective hybrids. A total of 21,392 germplasm accessions, including 750 accessions of wild species of genera *Pennisetum* and *Cenchrus*, collected from 50 countries are conserved at the ICRISAT Genbank, while 8,284 accessions are conserved at the National Bureau of Plant Genetic Resources (NBPGR), New Delhi, India. These lines have been widely used in breeding programs in both the public and private sectors for the development and commercialization of a large number of hybrids (public 70 and private sectors 105 were under cultivation in 2019).

Various molecular markers and genomic tools have been developed and applied for QTLs/genes identification, genetic diversity, and MAB to enhance pearl millet breeding by exploring its genetic potential at the molecular level (Serba and Yadav, 2016; Anuradha et al., 2017; Bollam et al., 2018; Kumar et al., 2018; Singhal et al., 2018; Ambawat et al., 2020; Srivastava et al., 2020a,b). Various molecular markers developed for pearl millet include restriction fragment length polymorphism [RFLP (Liu et al., 1994)], random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism [AFLP (Devos et al., 1995)], sequence-tagged sites [STSs (Allouis et al., 2001)], simple sequence repeat [SSRs (Qi et al., 2004; Meena et al., 2020; Srivastava et al., 2020a)], single-stranded conformation polymorphism-SNP [SSCP-SNP (Bertin et al., 2005)], expressed sequence tag-derived simple sequence repeats [EST-SSRs (Senthilvel et al., 2008; Rajaram et al., 2013)], DArT array technology [DArTs (Senthilvel et al., 2010; Supriya et al., 2011)], conserved intron-specific primers [CISP (Sehgal et al.,

TABLE 1 | Achievements and milestones in pearl millet improvement through conventional/heterosis breeding.

Year	Achievements	References
1940	Indian Council of Agricultural Research started Pearl millet breeding in India.	Yadav and Rai, 2013; Singh et al., 2014
1950	X1 and X2 were the two chance hybrids released in India for commercial utilization.	Yadav and Rai, 2013; Singh et al., 2014
1965	First pearl millet hybrid, Hybrid "HB-1" (Hybrid Bajra-1) was released.	Athwal, 1965
1965	Establishment of the ICAR-All India Coordinated Research Project on Pearl Millet	Yadav and Rai, 2013
1997–2008	Effective phenotypic screening techniques and resources were developed, and resistance breeding programs were developed.	Thakur and King, 1988; Singh et al., 1997; Hash et al., 1999; Hash and Witcombe, 2002; Jones et al., 2002; Thakur et al., 2008
1974–2020	Identification and release of 175 hybrids and 62 varieties for cultivation in different agro-ecological zones of India	Satyavathi et al., 2020
2018	Identified heterotic groups for grain yield and a total of 343 hybrid parental [maintainer (B-) and restorer (R-)] lines were genotyped with 88 polymorphic SSR markers.	Ramya et al., 2018
2021	Genetic variation studies for NUE among a set of 380 diverse pearl millet lines	Pujarula et al., 2021
2021	Studied expression of heat stress-responsive gene Pgghsp	Sankar et al., 2021

2012)], and single nucleotide polymorphisms [SNPs (Sehgal et al., 2012)]. High allelic variation and polymorphism were recorded between pairs of parental lines of pearl millet after screening with SNP markers, and they were also mapped on all the seven linkage groups reflecting the distribution of the markers in the genome (Sehgal et al., 2012). Similarly, a panel 21,663 SNP markers were also discovered, depicting > 5% of minor allele frequencies by Diack et al. (2020). Different types of molecular markers and genomic approaches ultimately provide a systematic knowledge on plant biology enabling MAB, which can accelerate the process of the development of new and resistant hybrids/varieties.

MAS and gene introgression into desirable genetic background have been proved to be very efficient for crops improvement as it reduces the cumbersome procedure of phenotypic evaluation and selection. Pearl millet is one of the crops where MAB approach has been applied to develop downy mildew-resistant variety "Improved HHB 67" (Hash et al., 2006). Later, studies were also conducted to identify and map major QTLs affecting abiotic stress tolerance in pearl millet (Yadav et al., 1999, 2002, 2004, 2011; Serraj et al., 2005; Bidinger et al., 2007; Kholová et al., 2012, 2013; Tharanya et al., 2018). The first

TABLE 2 | Achievements and milestones in pearl millet improvement through molecular and advanced genomic tools.

Year	Achievements	References
Molecular markers		
1994	Study of 200 samples of varied pearl millet lines using RFLP markers to reveal polymorphism	Liu et al., 1994
1995	Development of amplified fragment length polymorphism (AFLP) markers for pearl millet using nuclear genomic sequences	Devos et al., 1995
2000	163 AFLP markers were used to study genetic variability within and between pearl millet landraces	Busso et al., 2000
2001	Development of STSs markers from BAC clones	Allouis et al., 2001; Qi et al., 2001
2002	Genetic diversity was studied within and between 504 landraces of core collection using RFLP probes	Bhattacharjee et al., 2002
2003	18 SSR markers were developed from genomic sequences in pearl millet	Budak et al., 2003
2004	A consensus map of 353 RFLP and 65 SSR markers was developed for the first time.	Qi et al., 2004
2005	SSCP-SNP primes were developed from pearl millet EST collections	Bertin et al., 2005
2008	EST-based SSRs were developed in pearl millet	Senthilvel et al., 2008
2011	DArT platform was established for pearl millet, and 574 polymorphic DArT markers were mapped and used to genotype a set of 24 diverse pearl millet inbred lines	Supriya et al., 2011
2012	Conserved intron-specific primers (CISP) were developed from EST sequences using parents of two mapping populations for 18 genes	Sehgal et al., 2012
2013	Consensus linkage maps based on SSRs were constructed using four RIL populations	Rajaram et al., 2013
2013	Genetic diversity was analyzed in a novel set of restorer lines using SSR markers in pearl millet	Satyavathi et al., 2013
2014	Development of ISSR-based SCAR markers in pearl millet	Jogaiah et al., 2014
2015	Identification of single nucleotide polymorphisms (SNPs) using GBS platform	Hu et al., 2015; Sehgal et al., 2015
2020	A panel of 21,663 SNP markers was developed	Diack et al., 2020
2020	Morphological and molecular genetic diversity analysis of pearl millet (<i>Pennisetum glaucum</i>) maintainers and restorers.	Chandra et al., 2020
QTL mapping		
1995–2007	DNA markers have been established for around 60 different putative DM resistance QTLs in pearl millet	Jones et al., 1995, 2002; Hash et al., 1999; Breese et al., 2002; Hash and Witcombe, 2002; Gulia et al., 2007

(Continued)

TABLE 2 | Continued

Year	Achievements	References
1998	Identified molecular markers for three rust loci and one <i>Magnaporthe</i> resistance locus in pearl millet.	Morgan et al., 1998
1999, 2002	QTL was identified for grain and stover yield in pearl millet under terminal drought stress conditions	Yadav et al., 1999, 2002
2005	Evaluated putative DT-QTL on LG2	Bidinger et al., 2005; Serraj et al., 2005
2007	Three major QTLs were identified on LG 2, LG 3, and LG 4 for grain yield under variable post-flowering water conditions	Bidinger et al., 2007
2010	Identified QTL responsible for terminal drought tolerance	Kholová et al., 2010
2011	Identified one major QTL on LG2 for grain yield and drought tolerance	Yadav et al., 2011
2012	Co-mapped alleles to terminal drought tolerance QTL	Kholová et al., 2012
	Investigated the effects of DT-QTL on LG 2 under salinity stress	Sharma et al., 2014
2015	Detected four QTLs linked to high transpiration rate	Aparna et al., 2015
2016	A major QTL for rust resistance was identified on LG 1 using the cross 81B-P6 × ICMP 451-P8	Ambawat et al., 2016
2018	QTLs were identified on LG 1 and LG 4 for downy mildew (DM) resistance in pearl millet	Taunk et al., 2018
2018	Identified five QTLs for early drought stress conditions and stay-green trait	Debieu et al., 2018
2018	Role of chitosan nanoparticles was explored for resistance against pearl millet downy mildew	Siddaiah et al., 2018
2019	Introgressed DT-QTLs into hybrid HHB 226 from 863 B, the male parent HBL 11	Jangra et al., 2019
2019	Five QTLs were identified for resistance to three different pathotype isolates of <i>S. graminicola</i>	Chelpuri et al., 2019
Advanced genomic tools		
2006	Improved HHB67 was developed using the marker-assisted selection	Hash et al., 2006
2009	GWAS approach was used to dissect complex traits in pearl millet	Saïdou et al., 2009
2015	SNP markers were identified using genotyping-by-sequencing (GBS) pearl millet and high-density maps were constructed.	Hu et al., 2015; Moumouni et al., 2015
2015	PMiGAP was established and used for fine mapping of a drought tolerance DT-QTL on LG2 using candidate gene-based association mapping (AM) approach.	Sehgal et al., 2015

(Continued)

TABLE 2 | Continued

Year	Achievements	References
2016	Transcriptomic analysis was done using NGS tool to understand the mechanisms underneath resistance to downy mildew in pearl millet	Kulkarni et al., 2016
2016	16,650 SNPs and 333,567 sequence tags across all seven chromosomes have been identified for leaf spot resistance using GBS platform.	Punnauri et al., 2016
2017	A draft genome sequence of <i>S. graminicola</i> pathotype 1 of 299,901, 251 bp in length, N ₅₀ of 17,909 bp with a minimum of 1 Kb scaffold size was assembled	Nayaka et al., 2017
2017	Whole-genome sequence of pearl millet was deciphered	Varshney et al., 2017
2018	A total of 392,493 SNPs identified using GWAS on a panel of 188 inbred lines	Debieu et al., 2018
2018	Sequencing data were generated for RAD-seq and tGBS using genomic selection (GS) schemes in pearl millet.	Liang et al., 2018
2018	Transcriptome analysis for drought stress response using RNA-Seq approach in pearl millet	Dudhate et al., 2018; Jaiswal et al., 2018
2019	Genome of <i>Magnaporthe grisea</i> strain PMg_DI and was sequenced, and 13.1-Gb PE reads were generated	Prakash et al., 2019
2019	Studied the genetic diversity, population structure, and linkage disequilibrium in 398 accessions using GBS	Serba et al., 2019
2019	Genetic diversity of 130 forage-type hybrid parents of pearl millet was investigated using GBS-derived 7870 SNPs	Ponnaiah et al., 2019
2020	Characterized 309 inbred lines by 54,770 GBS-SNPs and reported higher nucleotide diversity in the panel derived from landraces and improved varieties from Africa and India.	Kanfany et al., 2020
2020	Comparative transcriptomics at vegetative and flowering stage was done using RNAseq analysis in a drought-tolerant (PRLT2/89-33) genotype to discover underlying genes to drought tolerance	Shivhare et al., 2020
2020	6920 genes and 6484 genes differentially expressed under heat stress and drought stress were identified using RAD-GBS	Sun et al., 2020
2020	Heterotic groups were defined and 0.9 million SNPs clustered into 12 R- and 7 B-line groups	Gupta et al., 2020
2021	Importance of wild relatives of pearl millet germplasm was harnessed for germplasm enhancement and improving biotic stress tolerance in pearl millet	Sharma et al., 2021

(Continued)

TABLE 2 | Continued

Year	Achievements	References
Quality improvement		
2009–2020	Implementation of biofortification approaches in pearl millet for high Fe and Zn	Govindaraj et al., 2009, 2019, 2020; Kanatti et al., 2014; Rai et al., 2016; Anuradha et al., 2017
2016, 2017	Reported diversity in the rancidity profile of pearl millet genotypes	Datta Mazumdar et al., 2016; Goyal and Chugh, 2017
2016	Identified two QTLs for grain Fe content on LG3 and LG5, and two QTLs for grain Zn content on LG3 and LG7 using replicated samples of 106 pearl millet RILs (F ₆) derived from ICMB 841-P3 × 863B-P2 cross.	Kumar et al., 2016
2017	Identified favorable alleles for grain iron and zinc content through AM in pearl millet.	Anuradha et al., 2017
2018	Reported large effect of Fe and Zn content QTLs using DArT and SSR markers in ICMS 8511-S1-17-2-1-1-B-P03 × AIMP 92901-S1-183-2-2-B-08 cross.	Kumar et al., 2018
2019	RAD-GBS was used and three GS models were implemented and compared using grain yield and dense molecular marker information of pearl millet	Jarquín et al., 2019
2020	Developed a rancidity matrix (RM) and classified pearl millet genotypes into low, medium, and high rancid groups	Goswami et al., 2020; Kumar et al., 2020
2020	Identified candidate genes for grain Fe and Zn contents in pearl millet	Mahendrakar et al., 2020
2021	Mapped multi-environment QTLs for grain iron and zinc contents in pearl millet.	Singhal et al., 2021

QTL mapping study was conducted on drought tolerance using recombinant inbred line (RIL) mapping population from the cross H 77/833-2 × PRLT 2/89-33, and a major QTL for drought tolerance (DT-QTL) was mapped on LG 2 explaining 32% variation for grain yield (Yadav et al., 1999, 2002). Subsequently, fine mapping of the DT-QTL was done and several markers mapped on this region of LG2 were identified, which were found to be linked with traits of drought tolerance, including delayed leaf senescence and leaf rolling under drought stress (Sehgal et al., 2012). Subsequently, Sharma et al. (2014) investigated the effects of DT-QTL under salt stress conditions. Later, Sehgal et al. (2015) utilized PMiGAP for the first time for fine mapping of a drought tolerance (DT) QTL (localized on linkage group 2) using candidate gene-based AM approach. This PMiGAP can be used as a genetic resource for GWAS studies in pearl millet.

Recently, NGS and other advanced high-throughput assays were used to sequence the pearl millet genome, which will prove useful for its improvement and enhancing different yield and

yield-related traits along with major biotic and abiotic stress tolerance and nutritionally significant traits worldwide (Varshney et al., 2017). Consequently, GBS, which allows simultaneous SNP discovery and genotyping, has been also used extensively in pearl millet to characterize germplasm (Poland and Rife, 2012; Ramu et al., 2013; Sehgal et al., 2015; Ponnaiah et al., 2019). GBS was used to construct a high-density genetic map with a more uniform distribution of markers in pearl millet from a biparental population in comparison with maps constructed earlier (Moumouni et al., 2015). Similarly, 83,875 SNP markers were identified from *PstI*-*MspI* reduced representative libraries of pearl millet by GBS, revealing a wide genetic variation in germplasm collection (Hu et al., 2015). A total of 398 accessions were used to study the genetic diversity, population structure, and linkage disequilibrium using GBS (Serba et al., 2019). Recently, Kanfany et al. (2020) characterized 309 inbred lines derived from landraces and improved varieties of India and Africa by 54,770 GBS-SNPs and reported higher nucleotide diversity in the panel.

Various genetic and genomic resources like GWAS and genomic selection (GS) have been also established in pearl millet, which is useful in trait discovery and whole-genome scan studies. Phenotypic data are combined with GBS data, and different genomic regions governing traits of interest are identified using GWAS. It can be used to exploit the natural diversity available in population or germplasm panels and can be used to increase mapping resolution in comparison with linkage mapping populations (Srivastava et al., 2020a). GS is eventually the extension of marker-assisted selection, and using it, genome-wide molecular markers are targeted and promising individuals are selected. It calculates the breeding values by combining marker genotypic data with phenotypic data collected from several environments and pedigree or kinship for improvement in quantitative traits and can ultimately speed up the genetic gain (Meuwissen et al., 2001; Goddard and Hayes, 2007; Heffner et al., 2009). Srivastava et al. (2020a) reviewed various case studies and the development of various whole-genome prediction/GS models based on GWAS which will be highly useful to explore the underlying genetics associated with pearl millet. The use of GWAS approach to dissect complex traits was initially reported in pearl millet by Saïdou et al. (2009). In this study, they revealed the genetic factors controlling the variations in flowering time at the phytochrome C (PHYC) (866 bp) locus, which plays a major role in crop adaptation mechanism. Later, Saïdou et al. (2014) investigated an extra region of 100 Kb around the *PHYC* gene using the same panel of 90 inbred lines to identify tightly linked best candidate markers. Debieu et al. (2018) also performed GWAS on a panel of 188 inbred lines of West Africa to identify QTLs associated with stay-green trait and biomass production in early drought stress conditions. In addition, different “omics” technologies such as transcriptomics, proteomics, and metabolomics can be also useful for quantitative and qualitative analyses of gene expression allowing more precise use of MAS and transgenic technologies. RNA sequencing (RNA-Seq) is widely used among the different transcriptome analysis methods as it can efficiently detect unknown genes and novel transcripts and has much potential to study gene expression

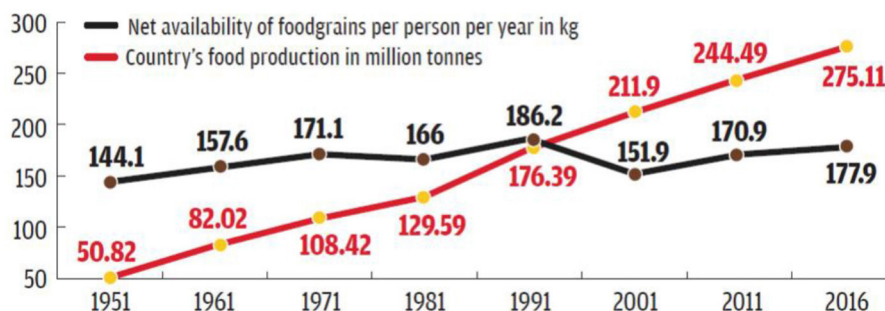


FIGURE 3 | Trends in food grain production and per capita availability in India. Adapted from: Agriculture Statistic, Ministry of Agriculture and Family Welfare.

and their regulating pathways (Hrdlickova et al., 2017). The transcriptome analysis of pearl millet can reflect new prospects into gene regulatory networks existing in this crop under abiotic stress conditions. Stress-regulated pathways in pearl millet can be studied by the detection and characterization of stress-responsive genes *via* transcriptomics and then different approaches can be followed for improving stress tolerance/resistance in millet (Mishra et al., 2007). Several transcriptomic studies in pearl millet have been used to reveal the functions of some salinity stress-responsive genes such as PgDHN, PgDREB2A, PgVDAC, PgNHX1 (Desai et al., 2006; Mishra et al., 2007; Verma et al., 2007; Agarwal et al., 2010; Reddy et al., 2012; Singh et al., 2015). Recently, complete transcriptome analysis has been done in pearl millet for drought stress response (Dudhate et al., 2018; Jaiswal et al., 2018; Shivhare et al., 2020). Dudhate et al. (2018) unraveled the molecular mechanism governing drought tolerance in two pearl millet inbred lines, ICMB 843 and ICMB 863, using RNA-Seq approach, and it is the first report of the study of drought tolerance by RNA sequencing in pearl millet. More recently, an important study on RNAseq analysis was performed to assess the comparative transcriptomics at the vegetative and flowering stage in a drought-tolerant (PRLT2/89-33) genotype to discover underlying genes to drought tolerance (Shivhare et al., 2020).

Proteomics is another important technology to get information on protein concentrations, post-translational modifications (PTMs), protein–protein interaction, structures linked with stress tolerance, regulatory functions of proteins encoded by genes (Ghatak et al., 2017). Identification and characterization of stress-responsive genes and their proteins from pearl millet can help in defining stress-regulated pathways. Further, it can help design strategies to improve stress tolerance/resistance of pearl millet as well as other related crop plants. Several proteomic studies have been also carried out in pearl millet, which can provide a framework to investigate C_4 photosynthesis in pearl millet in more depth (Ghatak et al., 2016, 2021; Weckwerth et al., 2020). A shotgun proteomics approach (GEL-LC-Orbitrap-MS) was used by Ghatak et al. (2016) to identify 2,281 proteins from different tissues of pearl millet (root, seed, and leaf) showing significant changes under drought stress condition. Thus, a lot of information has been generated for the improvement of this crop but still, there is a big challenge

to identify various crucial genes responsible for its adaption to survival under different abiotic stresses conditions. In addition, the identification of genomic regions governing NUE and its use in pearl millet breeding programs *via* MAS can be also exploited for survival under adverse conditions. Moreover, exploring several new and advanced genomic tools will be also beneficial for the advancement of this crop to harness its suitability to adverse conditions and utilize its inbuilt capacity to ensure global food and nutritional security.

NUTRITIONAL SECURITY

Food grain production has increased but still, to feed the growing population and meeting out good health of the people in the present situation, nutritional security is very important (Figure 3).

Intake of all essential macronutrients and micronutrients (vitamins and minerals) through a balanced and diversified diet in sufficient quantities is very vital for an active and healthy life. Nutritional insecurity is a major challenge for the world's growing population, which mainly depends on a micronutrient-deficient cereal-based diet. Pearl millet is the primary source of energy for the semi-arid tropics and drought-prone regions of Asia and Africa after the major cereals such as wheat, rice, maize and sorghum due to its nutritionally superior grain enriched with high amounts of essential amino acids, proteins, better fat digestibility, vitamins and minerals (Table 3). It is a good source of carbohydrate, energy, RS, 92.5% dry matter, fat (5–7%), ash (2.1%), dietary fiber (1.2/100g), 13.6% crude protein, quality protein (8–19%), 63.2% starch, α -amylase activity, minerals (2.3 mg/100g), vitamins A and B, and antioxidants such as coumaric acids and ferulic acid (Goswami et al., 2020). In addition, it is rich in unsaturated fatty acids (75%) and phytic acid, which are considered to be useful in lowering cholesterol and phytate, which in turn reduces cancer risk. It also exhibits antioxidant activities due to the presence of polyphenols, anthocyanins, phytates, phytosterols, tannins and pinacosanols and thus plays a significant role in aging. It is also enriched with many essential amino acids except lysine and threonine and has relatively higher methionine. Being gluten-free, it is extremely useful for people suffering from celiac diseases who are generally allergic to the gluten content of wheat and other cereals. Pearl

TABLE 3 | Nutritional comparison of pearl millet with sorghum, rice, and wheat (in 100 g grain).

Contents	Crop			
	Pearl millet	Sorghum	Rice	Wheat
Carbohydrates (g)	61.8	67.7	78.2	64.7
Protein (g)	10.9	09.9	07.9	10.6
Fat (g)	5.43	1.73	0.52	1.47
Energy (Kcal)	347	334	356	321
Dietary fiber (g)	11.5	10.2	02.8	11.2
Calcium (mg)	27.4	27.6	07.5	39.4
Phosphorus (mg)	289	274	96	315
Magnesium (mg)	124	133	19	125
Zinc (mg)	2.7	1.9	1.2	2.8
Fe (mg)	6.4	3.9	0.6	3.9
Thiamine (mg)	0.25	0.35	0.05	0.46
Riboflavin (mg)	0.20	0.14	0.05	0.15
Niacin (mg)	0.9	2.1	1.7	2.7
Folic acid (μ g)	36.1	39.4	9.32	30.1

Adapted from: NIN, Hyderabad, 2018.

millet is exceptionally useful for people suffering from diseases like diabetes, obesity, diabetic heart disease, atherosclerosis and metabolic diseases due to its health beneficial properties (Kumar et al., 2020).

It is also called the “Powerhouse of Nutrition” due to its richness with essential nutrients in good quantity and quality, which are vital for leading healthy and nutritious life. Pearl millet has elevated contents of various macronutrients as well as micronutrients like iron, zinc, magnesium, calcium, phosphorous, copper, manganese, riboflavin, and folic acid. Owing to such excellent nutritional values, it is gaining popularity and is preferred by people all over the world including developed countries.

Despite its nutritional superiority, the consumption of pearl millet flour is restricted to very few specific regions of the world because of the poor shelf life of the flour and the development of rancidity or off-odor on storage (Rani et al., 2018). Rancidity is caused by oxidative/hydrolytic enzymes such as lipase, lipoxygenase (LOX), etc., where they hydrolyze the triacylglycerol (TAG) to diacylglycerols, glycerol, monoglycerol, and free fatty acids (Manley and Mayer, 2012) (Figure 4).

Goswami et al. (2020) categorized 93 diverse genotypes of pearl millet into low, medium and high rancid groups. They developed and validated a rancidity matrix (RM) having three groups and six classes, which can prove very useful for the pearl millet breeders to develop low rancid pearl millet lines (Figure 5).

Pearl millet is used in different forms at a global level: unleavened bread (roti or chapatti), porridge, gruel, dessert etc. and it is generally defined as a poor man’s bread. Its flour can substitute (10–20%) for wheat flour in “whole-grain” breads, pretzels, crackers, tortillas and dry and creamed cereals (Dahlberg et al., 2003). In India and Africa, it is primarily grown for food and forage, while in the American continents, it is a main component of poultry and livestock sector (Serba et al., 2020).

Value addition is also very useful to promote its consumption, and several value-added products are gaining popularity among people. It can be used for making various traditional food products such as khichri, roti, sakarpore, gulgule etc., while industries are also using it for making products such as noodles, pasta, vermicelli, biscuits, bread, cookies, cakes, puffs etc in India (Figure 6). Several indigenous foods and drinks are made from flour/meal and malt of the millet in Africa and are nutritionally superior to other cereals. They contain high protein (up to 9.5/100 g for teff and fonio), ash, calcium (up to 344 mg/100 g for finger millet), phosphorus and potassium (up to 250 mg/100 g), iron and zinc levels (Obilana and Manyasa, 2002). The main food items prepared from pearl millet vary in the different countries of West Africa. The stiff or thick porridges (Tuwo or Tô) are very famous and generally used in all the Sahelian countries, while steam-cooked product “Couscous” is more common in the Francophone countries, including Senegal, Mali, Guinea, Burkina Faso, Niger, and Chad. The thin porridge “bouillie” is also popular in these countries. In Nigeria and Niger, the thin porridge “Fourra” and *Masa*, a fried cake are very popular, while “Soungouf,” “Sankhal” and “Araw” are mainly popular in Senegal (Kaur et al., 2014; Ajeigbe et al., 2020). Its grains are also locally brewed to produce non-alcoholic or alcoholic beverages in Asia and Africa (Dwivedi et al., 2012).

Among the 26 major risk factors for the disease estimates at the global level, Fe deficiency positions at ninth and Zn deficiency falls at 11th position (Ezzati et al., 2002). This issue is mainly severe in developing countries and infants, pregnant women, and adolescent children are particularly among high-risk groups. In India, ~80% pregnant women, 52% non-pregnant women and 74% of children lying in the age group of 6–35 months are suffering from anemia caused by Fe deficiency (Kramer and Allen, 2015). On the other hand, 50% of the world population is affected by zinc (Zn) deficiency, ultimately leading to diarrhea, impaired physical growth, and suppressed immune system (Gibson et al., 2008). In this context, biofortification of staple crops is a good, cost-effective, and sustainable approach to address malnutrition caused by micronutrients (Stein et al., 2007; Bouis et al., 2011). Biofortification is a multidisciplinary approach that can utilize the full potential of crop improvement and nutrition science and eliminate the prolonged issue of micronutrient malnutrition. Recent study showed that consuming biofortified pearl millet at 250 g day⁻¹ can meet 84% of the RDA for iron and 100% of the RDA for zinc in non-pregnant non-lactating (NPNL) women, while ordinary pearl millet can fulfill only 20% of their iron requirement (Neeraja et al., 2017). Absorption or bioavailability of iron in pearl millet is enough to meet out >50% of the daily requirement of adult males or children. In pearl millet, the bioavailability of Fe can be assumed to be 7.0–7.5% (Govindaraj et al., 2019). Approximately 50–100% of the daily requirement of iron can be met out with one meal of biofortified high-iron variety and is sufficient enough to overcome deficiency of iron in children, women and men (Kodkany et al., 2013; Mahendrakar et al., 2019).

Pearl millet as such is already a high-iron crop along with high Zn content; however, commercially available cultivars have lower Fe and Zn contents. Biofortification priority index (BPI)

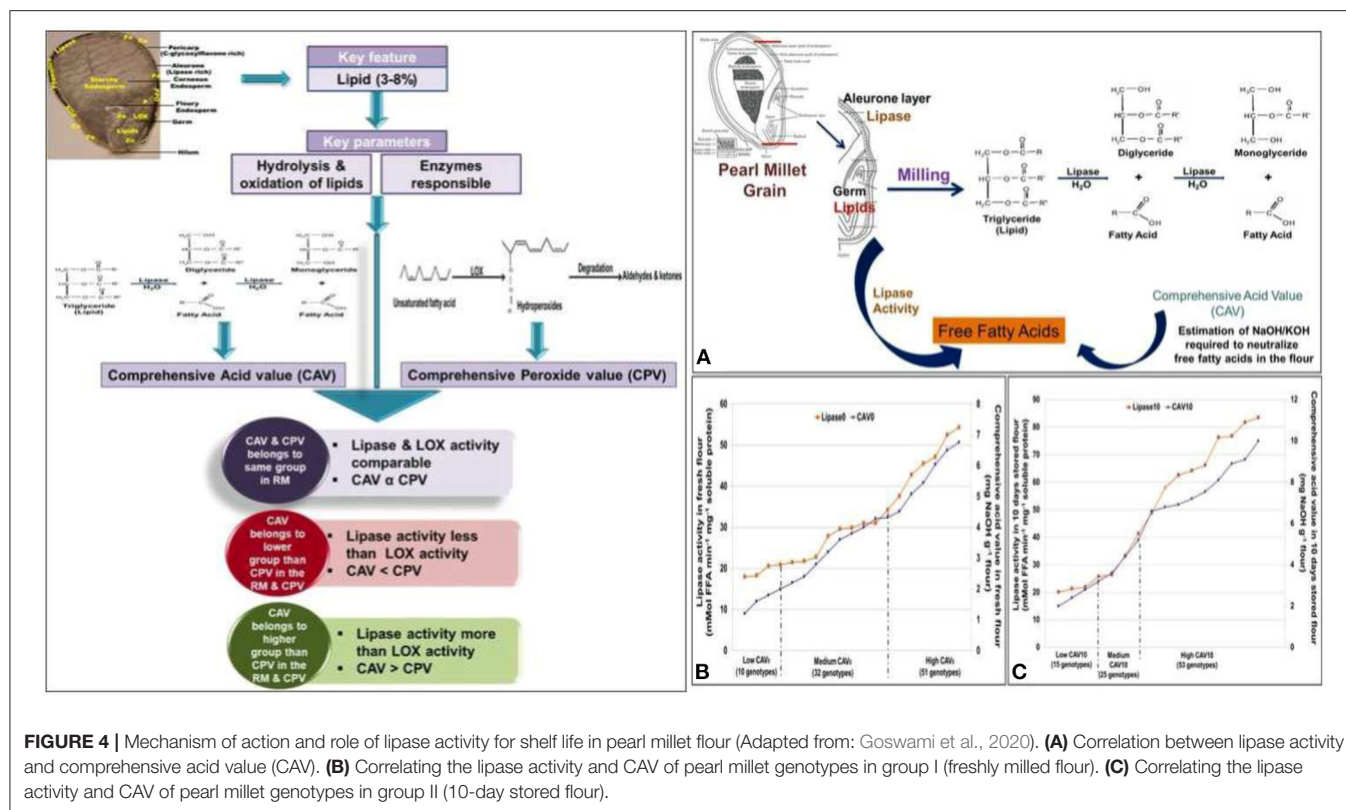


FIGURE 4 | Mechanism of action and role of lipase activity for shelf life in pearl millet flour (Adapted from: Goswami et al., 2020). **(A)** Correlation between lipase activity and comprehensive acid value (CAV). **(B)** Correlating the lipase activity and CAV of pearl millet genotypes in group I (freshly milled flour). **(C)** Correlating the lipase activity and CAV of pearl millet genotypes in group II (10-day stored flour).

indicates that pearl millet is a major target crop for iron and zinc biofortification. Enormous genetic variability (30–140 mg/kg Fe and 20–90 mg/kg Zn) can be used efficiently for developing high-yielding cultivars along with high iron and zinc densities (Kanatti et al., 2014). Targeting the development of biofortified crops with enhanced nutrients, ICAR is supporting a Consortia Research Platform (CRP) of rice, wheat, maize, sorghum, pearl millet and small millets since 2014. Several promising donors have been identified and breeding material is being generated combining high nutrient content and yield. Thus, focus was also laid on nutritional improvement in addition to yield improvement in pearl millet. Development of Fe and Zn-biofortified varieties/hybrids was considered as a high-priority area in pearl millet resulting in the inclusion of biofortification in the main stream and now it is a routine affair. The gradual shift of efforts of biofortification research from exploring variability in 2008 to breeding high-Fe cultivars by exploiting existing high-Fe lines in 2011 and testing and delivery of biofortified varieties and hybrids was recognized in India and became conscious breeding part in pearl millet at various centers of public and private partners. The first high-iron pearl millet variety “Dhanashakti” was released for the state of Maharashtra in 2013, and subsequently, it was notified and released by Central Variety Releasing Committee in April 2014, for growing in all pearl millet-growing regions of India. Inclusion of nutritional quality traits for Fe (42 ppm) and Zn (32 ppm) in the varietal promotion criteria is first of its kind in any of the food crop and the world too (Satyavathi, 2017). Since then, many hybrids/varieties were

developed as micronutrient-rich pearl millet cultivars, including the seven biofortified cultivars identified from the special biofortification trial. In addition, a project on “Development of Biochemical and Physical Processing Technology to Arrest Oxidation of Lipids/Flavones to Enhance the Shelf-Life of Pearl Millet Flour” under Niche area of excellence (NAE) program has been started in ICAR-IARI, India, to overcome the issue of rancidity and enhance shelf life of pearl millet flour. Parallel research programs addressing the issues of rancidity were also taken up in ICRISAT along with ICAR.

Screening of enormous genetic materials such as germplasm collections, elite lines, hybrids, segregating populations etc. as well as phenotyping for iron and zinc through destructive techniques and breeding resources is required to breed micronutrient dense cultivars. Several studies have been conducted by several research groups around the world to assess the genetic diversity for Fe and Zn contents from time to time using different sets of germplasm accessions and breeding lines (Velu et al., 2007; Rai et al., 2012, 2014; Kanatti et al., 2014; Anuradha et al., 2017; Kumar et al., 2018; Govindaraj et al., 2019, 2020). Various studies were reported to show a significant importance of environment and $G \times E$ interactions to determine the levels of grain Fe and grain Zn contents in pearl millet. Such studies also identified donors that are high and stable micronutrient contents (Satyavathi et al., 2015; Rai et al., 2016; Pawar et al., 2018; Singhal et al., 2018). In addition, many efforts were put to discover, validate and deploy trait-based molecular markers for iron and zinc contents in pearl millet

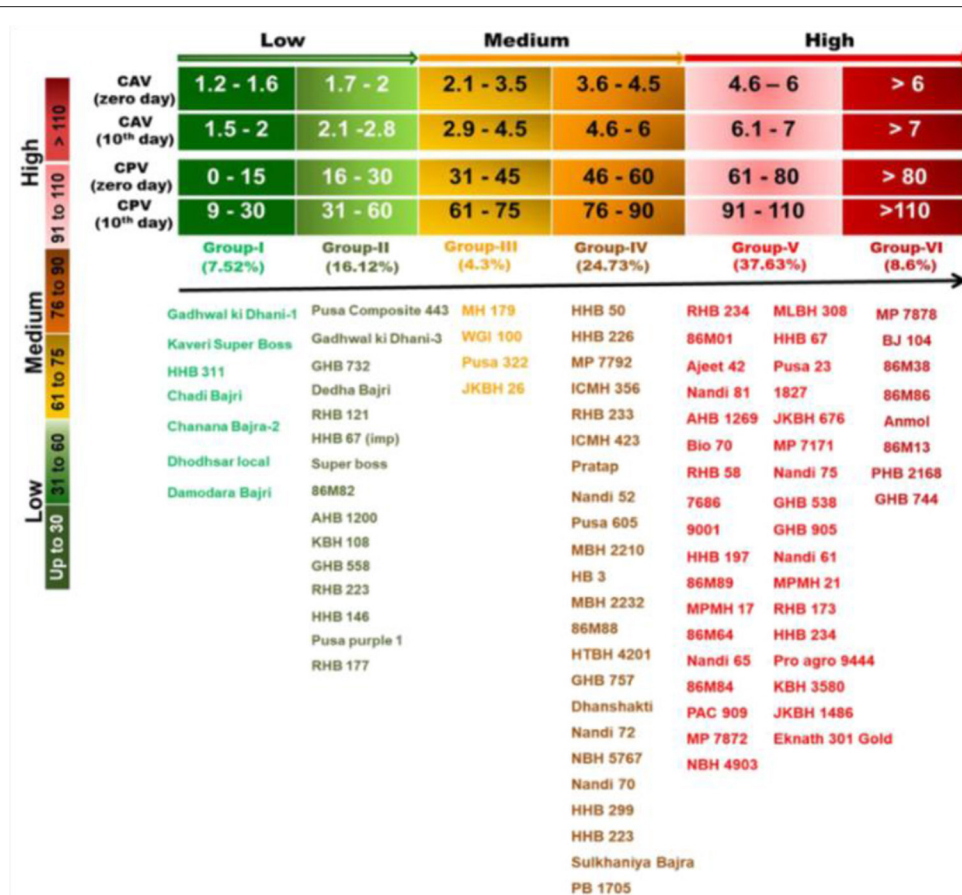


FIGURE 5 | Rancidity matrix to measure rancidity in the pearl millet flour (Adapted from: Goswami et al., 2020).

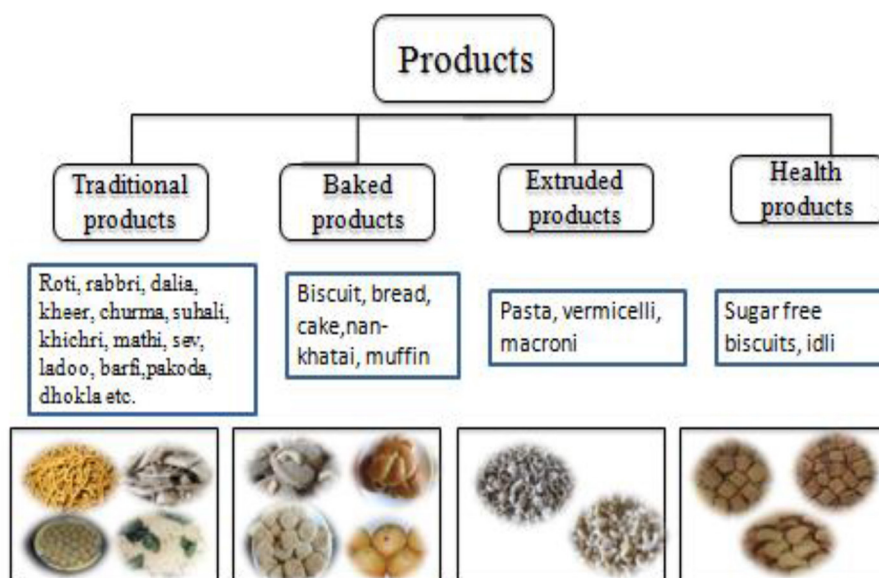
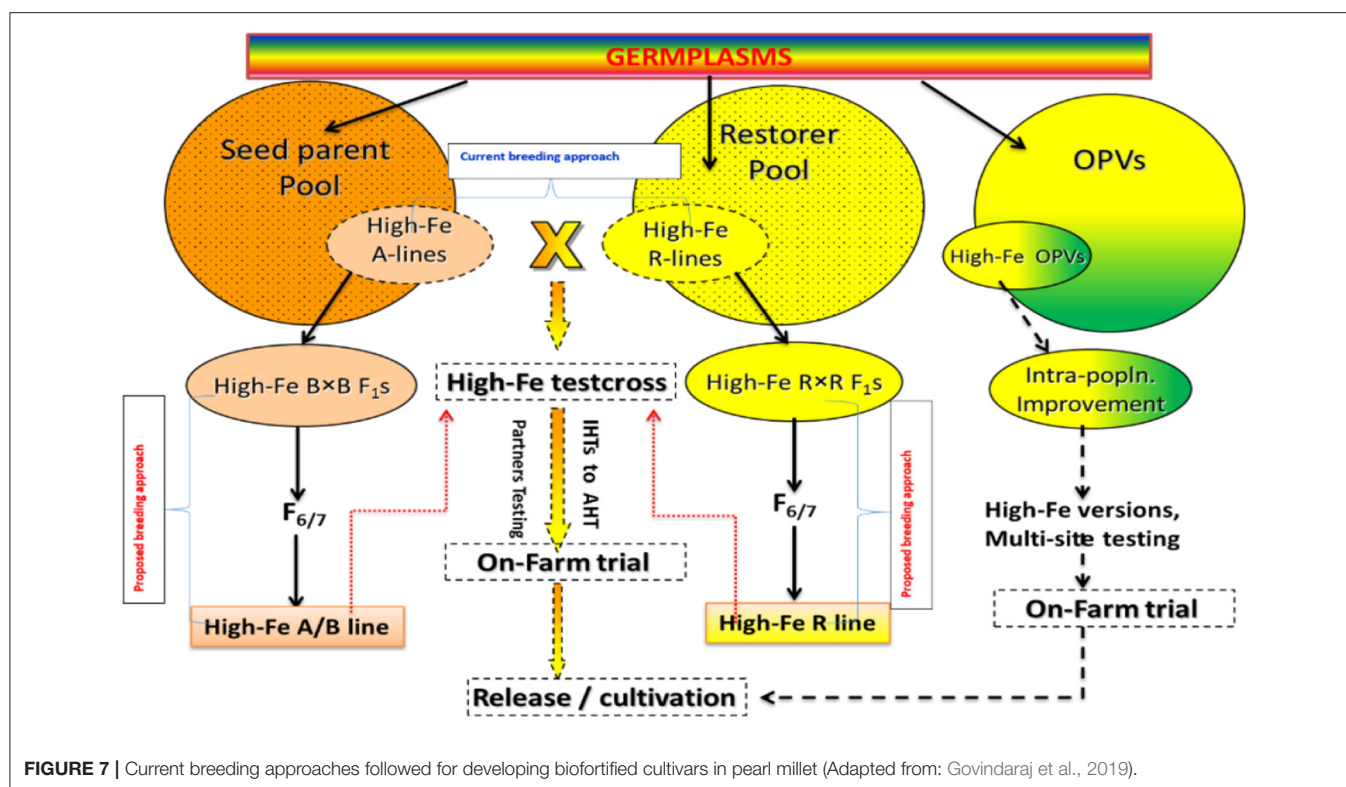


FIGURE 6 | Delicious and value-added food products from pearl millet.



(Kumar et al., 2016, 2018; Anuradha et al., 2017). Manwaring et al. (2016) elaborated the several challenges and opportunities associated with biofortification of pearl millet and emphasized that naturally occurring genetic variations existing in germplasm collections should be incorporated into elite, micronutrient-rich varieties through different advanced platforms to develop biofortified varieties.

The success of pearl millet biofortification program lies in the high-throughput precision phenotyping for the estimation of grain micronutrient content. Further, demonstration of the feasibility of developing high-Fe and high-yielding hybrids and encouraging the partners to breed for these micronutrient traits by means of mainstreaming the biofortification can be an important approach for its promotion. There is a high need to facilitate the focused screening of partners breeding materials for grain Fe and Zn contents for their introgression of high Fe content into locally adapted, high-yielding and farmer-preferred cultivars (Figure 7). Promotion and strengthening of the pipeline of high-iron and high-yielding partners-bred hybrids and testing of their grain samples for grain Fe and Zn contents can be a useful step for enhancing its importance. In addition to this, bioavailability is another aspect that needs to be focused to get full potential of this crop.

CHALLENGES FOR PEARL MILLET

Despite the breeding efforts, most of breeding programs fail to deliver hybrids due to a vast variation in microclimate (day and night temperature and humidity) and soil apart from rainfall, which requires proper quantification. Further, narrow cultivar

diversity in drought-prone ecology also is another factor for this. Thus, there is a high need to give higher priority to the below-mentioned areas to promote its production and utilization:

- Development of hybrids/varieties of pearl millet with better regenerative capacity on reversal of dry spell for harsh environment/drought-prone areas (for A₁ zone in India).
- Development of hybrids/varieties resistant/tolerant to salt/high temperature.
- Shift in focus of breeding from productivity improvement to the identification of end product-specific traits.
- Mainstreaming of biofortification in pearl millet for iron and zinc.
- Enhancement of shelf life of pearl millet flour and overcome rancidity to promote its products.
- Development of screening protocols and control measures against different diseases such as downy mildew, blast, rust, ergot, smut.
- Generating authentic data on nutritional benefits of pearl millet and bioavailability studies.
- A study on demand survey for pearl millet.

CONCLUSION

Pearl millet being a climate-resilient crop along with high nutritional value can be exploited for improving nutritional quality and combating malnutrition. It is almost free from major diseases and insect attacks and could be cultivated with a good harvest. Hence, the focus should be laid towards the development of food products from pearl millet to make

it acceptable as an alternative crop of the future. Indian policymakers need to refocus their interest toward millet farming systems and policies should be engraved for creating a feasible environment for pearl millet farmers. After the successful use of genomic tools, screening and development of improved genotypes have become easy and fast, and the progress toward enhancing the use of genomic resources is quite appreciable. But it can be characterized further for harnessing the natural genetic variation within the germplasm, and a lot of efforts are still required to execute genomics to improve the crop using high-throughput genome analysis, sequence-based molecular markers, NGS techniques and genome editing etc. The genomics and breeding platform needs a better alignment and constant up-gradation to develop improved hybrid parental lines, and populations must be adapted specifically according to the specific global agro-ecologies. The idea of genetic gains, genome editing, pre-breeding, and speed breeding can be also very useful for the researchers in selecting plants for desired traits along with many variations. On the other hand, the construction of high-density maps, QTL detection, candidate gene identification, new genome sequence techniques, use of advanced multi-parental and AM panels, GWAS and GS can speed up the recognition of allelic variants for pearl millet improvement. Mapping of several abiotic stresses, QTLs, etc. is highly desired to combat global climatic effects and the recently developed technologies must be tested under actual conditions.

The outcomes from model crops can be used in pearl millet to achieve added improvement and develop Zn- and Fe-enriched biofortified varieties. Synteny studies can prove useful for the identification of common genes linked with nutrition

biosynthesis pathways, and these should be incorporated into pearl millet by traditional breeding or transgene techniques for further nutritional improvements. In addition, nutritional as well as the economic security of small and marginal farmers, enhancing demand of pearl millet, value addition and market-led extension through food science and nutrition is vital to promote the cultivation and consumption of this crop. In conclusion, multidisciplinary approaches, including breeding, genomics, bioinformatics, biotechnology, nutrition and genetics etc. are required to exploit and harness the beneficial attributes of nutricereal pearl millet for combating changing climate and attaining nutritional security.

AUTHOR CONTRIBUTIONS

CTS has conceived the idea of writing this review, drafted and edited the manuscript. SA contributed in writing and editing the manuscript. RKS and VK contributed in manuscript preparation.

ACKNOWLEDGMENTS

We sincerely acknowledge Indian Council of Agricultural Research (ICAR), New Delhi for providing financial assistance under ICAR-All India Coordinated Research Project on Pearl millet, Jodhpur (AICRP on Pearl millet).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.659938/full#supplementary-material>

REFERENCES

- Agarwal, P., Agarwal, P. K., Joshi, A. J., Sopory, S. K., and Reddy, M. K. (2010). Overexpression of PgDREB2A transcription factor enhances abiotic stress tolerance and activates downstream stress-responsive genes. *Mol. Biol. Rep.* 37, 1125–1135. doi: 10.1007/s11033-009-9885-8
- Ajeigbe, H. A., Angarawai, I. I., Inuwa, A. H., Akinseye, F. M., and Abdul Azeez, T. (2020). *Hand Book on Improved Pearl Millet Production Practices in North Eastern Nigeria*, 9885–9888.
- Allouis, S., Qi, X., Lindup, S., Gale, M. D., and Devos, K. M. (2001). Construction of a BAC library of pearl millet, *Pennisetum glaucum*. *Theor. Appl. Genet.* 102, 1200–1205. doi: 10.1007/s001220100559
- Ambawat, S., Senthivel, S., Hash, C. T., Nepolean, T., Rajaram, V., Eshwar, K., et al. (2016). QTL mapping for rust resistance in pearl millet using an integrated DArT and SSR-based linkage map. *Euphytica*. 209, 461–476. doi: 10.1007/s10681-016-1671-9
- Ambawat, S., Singh, S., Shobhit, Meena R. C., and Satyavathi, C.T. (2020). “Biotechnological applications for improvement of the pearl millet crop.” in: *Pearl Millet: Properties, Functionality and its Applications*, eds S. K. Gahlawat, S. Punia, A. K. Siroha, K. S. Sandhu, and M. Kaur (Boca Raton, FL: Taylorand Francis (CRC Press)), 115–138.
- Anuradha, N., Satyavathi, C.T., Meena, M. C., Sankar, S. M., Bharadwaj, C., Bhat, J., et al. (2017). Evaluation of pearl millet [*Pennisetum glaucum* (L.) R. Br.] for grain iron and zinc content in different agro climatic zones of India. *Indian J. Genet. Plant Breed.* 77, 65–73. doi: 10.5958/0975-6906.2017.00009.8
- Aparna, K., Nepolean, T., Srivastava, R. K., Kholová, J., Rajaram, V., Kumar, S., et al. (2015). Quantitative trait loci associated with constitutive traits control water use in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Plant Biol.* 17, 1073–1084. doi: 10.1111/plb.12343
- Athwal, D. S. (1965). Hybrid bajara-1 marks a new era. *Indian Farm.* 15,6–7.
- Bertin, I., Zhu, J. H., and Gale, M. D. (2005). SSCP-SNP in pearl millet-a new marker system for comparative genetics. *Theor. Appl. Genet.* 110, 1467–1472. doi: 10.1007/s00122-005-1981-0
- Bhattacharjee, R., Bramel, J., Hash, T., Kolesnikova-Allen, A., and Khairwal, S. (2002). Assessment of genetic diversity within and between pearl millet landraces. *Theor. Appl. Genet.* 105, 666–673. doi: 10.1007/s00122-002-0917-1
- Bidinger, F. R., Nepolean, T., Hash, C. T., Yadav, R. S., and Howarth, C. J. (2007). Quantitative trait loci for grain yield in pearl millet under variable post flowering moisture conditions. *Crop Sci.* 47, 969–980. doi: 10.2135/cropsci2006.07.0465
- Bidinger, F. R., Serraj, R., Rizvi, S. M. H., Howarth, C. J., Yadav, R. S., and Hash, C. T. (2005). Field evaluation of drought tolerance QTL effects on phenotype and adaptation in pearl millet [*Pennisetum glaucum* (L.) R. Br.] topcross hybrids. *Field Crops Res.* 94, 14–32. doi: 10.1016/j.fcr.2004.11.006
- Bollam, S., Pujarula, V., Srivastava, R. K., and Gupta, R. (2018). “Genomic approaches to enhance stress tolerance for productivity improvements in pearl millet: genomic approaches”. in *Biotechnologies of Crop Improvement*, Vol. 3, eds S. S. Gosal and S. H. Wani. (Basel: Springer International Publishing AG, part of Springer Nature), 239–264.
- Bouis, H. E., Hotz, C., Mc Clafferty, B., Meenakshim, J.V., and Pfeiffer, W. H. (2011). Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 32, S31–S40. doi: 10.1177/15648265110321S105
- Bray, E. A., Bailey-Serres, J., and Weretilnyk, E. (2000). “Responses to abiotic stresses.” in *Biochemistry and Molecular Biology of Plants*, eds W. Gruissem, B. Buchanan, R. Jones (Rockville, MD: American Society of Plant Physiologists), 1158–1249.
- Breese, W. A., Hash, C. T., Devos, K. M., and Howarth, C. J., Leslie, J. F. (2002). Pearl millet genomics and breeding for resistance to downy mildew. *Sorghum Millets Dis.* 243–246. doi: 10.1002/9780470384923.ch41

- Budak, H., Pedraza, F., Cregan, P. B., Baenziger, P. S., and Dweikat, I. (2003). Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of pearl millet germplasm. *Crop Sci.* 43:2284. doi: 10.2135/cropsci2003.2284
- Busso, C. S., Devos, K. M., Ross, G., Mortimore, M., Adams, W. M., Ambrose, M. J., et al. (2000). Genetic diversity within and among landraces of pearl millet (*Pennisetum glaucum*) under farmer management in West Africa. *Genet. Res. Crop Evol.* 47, 561–8. doi: 10.1023/A:1008767220320
- Chandra, S., Singh, S. P., Kaswan, V., Chaudhary, S., Singh, A. M., Sankar, M., et al. (2020). Morphological and molecular genetic diversity analysis of pearl millet (*Pennisetum glaucum*) maintainers and restorers. *Indian J. Agric. Sci.* 90, 2083–2089
- Chelpuri, D., Sharma, R., Durga, K. K., Katiyar, P., Mahendrakar, M. D., Singh, R. B., et al. (2019). Mapping quantitative trait loci (QTLs) associated with resistance to major pathotype-isolates of pearl millet downy mildew pathogen. *Eur. J. Plant Pathol.* 154, 983–994. doi: 10.1007/s10658-019-01718-x
- Choudhary, S., Guha, A., Kholova, J., Pandravada, A., Essina, C. D., Cooper, M., et al. (2020). Maize, sorghum, and pearl millet have highly contrasting species strategies to adapt to water stress and climate change-like conditions. *Plant Sci.* 295:110297. doi: 10.1016/j.plantsci.2019.110297
- Dahlberg, J. A., Wilson, J. P., and Snyder, T. (2003). "Sorghum and pearl millet: health foods and industrial products in developed countries." in *Alternative Uses of Sorghum and Pearl Millet in Asia. Proceedings of Expert Meeting* (Patancheru: ICRIASAT). 42–59.
- Datta Mazumdar, S. K., Gupta, R., Banerjee, S., Gite, P., and Durgalla, P. B. (2016). *Determination of Variability in Rancidity Profile of Select Commercial Pearl Millet Varieties/Hybrids*. DC 24. Poster presented in CGIAR Research Program on Dryland Cereals Review Meeting held at Hyderabad, India, 5–6 October 2016. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Telangana, India. Available online at: <http://drylandcereals.cgiar.org/index.php>
- Debieu, M., Sine, B., Passot, S., Grondin, A., Akata, E., Gangashetty, P., et al. (2018). Response to early drought stress and identification of QTLs controlling biomass production under drought in pearl millet. *PLoS ONE*. 13:e0201635. doi: 10.1371/journal.pone.0201635
- Desai, M. K., Mishra, R. N., Verma, D., Nair, S., Sopory, S. K., and Reddy, M. K. (2006). Structural and functional analysis of a salt stress inducible gene encoding voltage dependent anion channel (VDAC) from pearl millet (*Pennisetum glaucum*). *Plant Physiol. Biochem.* 44, 483–493. doi: 10.1016/j.plaphy.2006.08.008
- Devos, K. M., Pittaway, T. S., Busso, C. S., Gale, M. D., Witcombe, J. R., and Hash, C. T. (1995). Molecular tools for the pearl millet nuclear genome. *Int. Sorghum Millets Newsl.* 36, 64–66.
- Diack, O., Kanfany, G., Gueye, M. C., Sy, O., Fofana, A., Tall, H. et al. (2020). GWAS unveils features between early- and late-flowering pearl millets. *BMC Genomics*. 21:777. doi: 10.1186/s12864-020-07198-2
- Directorate of Millets Development (2020). *Department of Agriculture, Co-operation & Farmers Welfare, Ministry of Agriculture & Farmers Welfare, Government of India*.
- Dita, M. A., Rispail, N., Prats, E., Rubiales, D., and Singh, K. B. (2006). Biotechnology approaches to overcome biotic and abiotic stress constraints in legumes. *Euphytica*. 147, 1–24. doi: 10.1007/s10681-006-6156-9
- Dudhate, A., Shinde, H., Tsugama, D., Liu, S., and Takano, T. (2018). Transcriptomic analysis reveals the differentially expressed genes and pathways involved in drought tolerance in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *PLoS ONE*. 13:e0195908. doi: 10.1371/journal.pone.0195908
- Dwivedi, S., Upadhyaya, H., Senthilvel, S., and Hash, C. (2012). "Millets: genetic, and genomic resources," in *Plant Breeding Reviews*, ed. J. Janick (Hoboken, NJ: John Wiley, and Sons, Inc.), 247–375.
- Dyall, S. C. (2015). Long-chain omega-3 fatty acids and the brain: a review of the independent and shared effects of EPA, DPA and DHA. *Front. Aging Neurosci.* 7:52. doi: 10.3389/fnagi.2015.00052
- Ezzati, M., Lopez, A. D., Rodgers, A., Vanderhoorn, S., and Murray, C. J. L. (2002). Selected major risk factors and global and regional burden of disease. *Lancet*. 360, 1347–1360. doi: 10.1016/S0140-6736(02)11403-6
- FAO (2019). *Food and Agriculture Organization of the United Nations*. Rome.
- Ghatak, A., Chaturvedi, P., Bachmann, G., Valledor, L., Ramšak, Ž., Bazargani, M. M., et al. (2021). Physiological and proteomic signatures reveal mechanisms of superior drought resilience in pearl millet compared to wheat. *Front. Plant Sci.* 11:600278. doi: 10.3389/fpls.2020.600278
- Ghatak, A., Chaturvedi, P., Nagler, M., Roustan, V., Lyon, D., Bachmann, G., et al. (2016). Comprehensive tissue-specific proteome analysis of drought stress responses in *Pennisetum glaucum* (L.) R. Br. (Pearl millet). *J. Proteomics*. 143, 122–135. doi: 10.1016/j.jpro.2016.02.032
- Ghatak, A., Chaturvedi, P., and Weckwerth, W. (2017). Cereal crop proteomics: systemic analysis of crop drought stress responses towards marker-assisted selection breeding. *Front. Plant Sci.* 8:757. doi: 10.3389/fpls.2017.00757
- Gibson, R. S., Hess, S. Y., Hotz, C., and Brown, K. H. (2008). Indicators of zinc status at the population level: a review of the evidence. *Br. J. Nutr.* 99, S14–S23. doi: 10.1017/S0007114508006818
- Goddard, M. E., and Hayes, B. J. (2007). Genomic selection. *J. Anim. Breed. Genet.* 124, 323–330. doi: 10.1111/j.1439-0388.2007.00702.x
- Goswami, S., Asrani, P., Ansheef Ali, T. P., Kumar, R. D., Vinutha, T., Veda, K. et al. (2020). Rancidity matrix: development of biochemical indicators for analysing the keeping quality of pearl millet flour. *Food Anal. Meth.* 13, 2147–2164. doi: 10.1007/s12161-020-01831-2
- Govindaraj, M., Rai, K. N., Cherian, B., Pfeiffer, W. H., Kanatti, A., and Shivate, H. (2019). Breeding biofortified pearl millet varieties and hybrids to enhance millet markets for human nutrition. *Agriculture*. 9:106. doi: 10.3390/agriculture9050106
- Govindaraj, M., Selvi, B., and Rajarathinam, S. (2009). Correlation studies for grain yield components and nutritional quality traits in pearl millet (*Pennisetum glaucum* (L.) R. Br.) Germplasm. *World J. Agric. Sci.* 5, 686–689.
- Govindaraj, M., Shanmugasundaram, P., Sumathi, P., and Muthiah, A. R. (2010). Simple, rapid and cost effective screening method for drought resistant breeding in pearl millet. *Elec J Plant Breed.* 1, 590–599.
- Govindaraj, M., Yadav, O. P., Rajpurohit, B. S., Kanatti, A., Rai, K. N., and Dwivedi, S. L. (2020). Genetic variability, diversity and interrelationship for twelve grain minerals in 122 commercial pearl millet cultivars in India. *Agric. Res.* 9, 516–525. doi: 10.1007/s40003-020-00470-7
- Goyal, P., and Chugh, L. K. (2017). Shelf life determinants and enzyme activities of pearl millet: a comparison of changes in stored flour of hybrids, CMS lines, inbreds and composites. *J. Food Sci. Technol.* 54, 3161–3169. doi: 10.1007/s13197-017-2752-z
- Gulia, S. K., Hash, C. T., Thakur, R. P., Breese, W. A., and Sangwan, R. S. (2007). "Mapping new QTLs for downy mildew [*Sclerospora graminicola* (Sacc.) J. Schroet.] resistance in pearl millet [*Pennisetum glaucum* (L.) R. Br.]." in *Crop Production in Stress Environments - Genetic and Management Options*, eds D. S. Singh, V. S. Tomar, R. K. Behl, S. D. Upadhyaya, and M. S. Bhale (New Delhi: Agrobios International).
- Gupta, S. K., Rai, K. N., Singh, P., Ameta, V. L., Gupta, S. K., Jayalekha, A. K., et al. (2015). Seed set variability under high temperatures during flowering period in pearl millet [*Pennisetum glaucum* L. (R.) Br.]. *Field Crops Res.* 171, 41–53. doi: 10.1016/j.fcr.2014.11.005
- Gupta, S. K., Sudarshan Patil, K., Rathore, A., Yadav, D. V., Sharma, L. D., Mungra, K. D., et al. (2020). Identification of heterotic groups in South-Asian-bred hybrid parents of pearl millet. *Theor. Appl. Genet.* 133, 873–888. doi: 10.1007/s00122-019-03512-z
- Hash, C. T., Singh, S. D., Thakur, R. P., and Talukdar, B. S. (1999). "Breeding for disease resistance". in *Pearl Millet Breeding*, eds I. S. Khairwal, K. N. Rai, D. J. Andrews, G. Harinarayana (New Delhi: Oxford and IBH), 337–379.
- Hash, C. T., Thakur, R. P., Rao, V. P., and Raj, A. B. (2006). Evidence for enhanced resistance to diverse isolates of pearl millet downy mildew through gene pyramiding. *Int. Sorghum Millets Newsl.* 47, 134–138.
- Hash, C. T., and Witcombe, J. R. (2002). "Gene management and breeding for downy mildew resistance". in: *Sorghum and Millets Diseases*, ed J. E. Leslie (Ames, IA: Iowa State Press), 27–36.
- Heffner, E. L., Sorrells, M. E., and Jannink, J. L. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12. doi: 10.2135/cropsci2008.08.0512
- Hrdlickova, R., Toloue, M., and Tian, B. (2017). RNA-Seq methods for transcriptome analysis. *Wiley Inter. Discip. Rev. RNA*. 8:e1364. doi: 10.1002/wrna.1364
- Hu, Z., Mbaké, B., Perumal, R., Guéye, M. C., Sy, O., Bouchet, S., et al. (2015). Population genomics of pearl millet (*Pennisetum glaucum* (L.) R. Br.): comparative analysis of global accessions and Senegalese landraces. *BMC Genomics*. 16:1048. doi: 10.1186/s12864-015-2255-0

- Jaiswal, S., Antala, T. J., Mandavia, M. K., Chopra, M., Jasrotia, R. S., Tomar, R. S., et al. (2018). Transcriptomic signature of drought response in pearl millet (*Pennisetum glaucum* (L.) and development of web-genomic resources. *Sci. Rep.* 8, 1–16. doi: 10.1038/s41598-018-21560-1
- Jangra, S., Rani, A., Yadav, R. C., Yadav, N. R., and Yadav, D. (2019). Introgression of terminal drought stress tolerance in advance lines of popular pearl millet hybrid through molecular breeding. *Plant Physiol. Rep.* 24, 359–369. doi: 10.1007/s40502-019-00464-w
- Jarquin, D., Howard, R., Liang, Z., Gupta, S. K., Schnable, J. C., and Crossa, J. (2019). Enhancing hybrid prediction in pearl millet using genomic and/or multi-environment phenotypic information of inbreds. *Front. Genet.* 10:e01294. doi: 10.3389/fgene.2019.01294
- Jogaiah, S., Sharathchandra, R. G., and Raj, N. (2014). Development of SCAR marker associated with downy mildew disease resistance in pearl millet (*Pennisetum glaucum* L.). *Mol. Biol. Rep.* 41, 7815–7824. doi: 10.1007/s11033-014-3675-7
- Jones, E. S., Breese, W. A., Liu, C. J., Singh, S. D., Shaw, D. S., and Witcombe, J. R. (2002). Mapping quantitative trait loci for resistance to downy mildew in pearl millet: field and glasshouse screens detect the same QTL. *Crop Sci.* 42, 1316–1323. doi: 10.2135/cropsci2002.1316
- Jones, E. S., Liu, C. J., Gale, M. D., Hash, C. T., and Witcombe, J. R. (1995). Mapping quantitative trait loci for downy mildew resistance in pearl millet. *Theor. Appl. Genet.* 91, 448–456. doi: 10.1007/BF00222972
- Jukanti, A. K., Gowda, C. L. L., Rai, K. N., Manga, V. K., and Bhatt, R. K. (2016). Crops that feed the world 11. Pearl millet (*Pennisetum glaucum* L.): an important source of food security, nutrition and health in the arid and semi-arid tropics. *Food Sec.* 8, 307–329. doi: 10.1007/s12571-016-0557-y
- Kanatti, A., Rai, K. N., Radhika, K., and Govindaraj, M. (2016). Genetic architecture of open-pollinated varieties of pearl millet for grain iron and zinc densities. *Indian J. Genet.* 76, 299–303. doi: 10.5958/0975-6906.2016.00045.6
- Kanatti, A., Rai, K. N., Radhika, K., Govindaraj, M., Sahrawat, K. L., Srinivasu, K., et al. (2014). Relationship of grain iron and zinc content with grain yield in pearl millet hybrids. *Crop Improv.* 41, 91–96.
- Kanfany, G., Serba, D. D., Rhodes, D., Amand, P., Bernardo, A., and Gangashetty, P. (2020). Genomic diversity in pearl millet inbred lines derived from landraces and improved varieties. *BMC Genomics.* 21:469. doi: 10.1186/s12864-020-06796-4
- Kaur, K. D., Jha, A., Sabikhi, L., and Singh, A. K. (2014). Significance of coarsecereals in health and nutrition: a review. *J. Food Sci. Technol.* 51, 1429–1441. doi: 10.1007/s13197-011-0612-9
- Kholová, J., Hash, C. T., Kakker, A., Kočová, M., and Vadez, V. (2010). Constitutive water-conserving mechanisms are correlated with the terminal drought tolerance of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *J. Exp. Bot.* 61, 369–377. doi: 10.1093/jxb/er314
- Kholová, J., McLean, G., Vadez, V., Craufurd, P., and Hammer, G. L. (2013). Drought stress characterization of post-rainy season (rabi) sorghum in India. *Field Crops Res.* 141, 38–46.
- Kholová, J., Nepolean, T., Hash, C. T., Supriya, A., Rajaram, V., Senthilvel, S., et al. (2012). Water saving traits co-map with a major terminal drought tolerance quantitative trait locus in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Mol. Breed.* 30, 1337–1353. doi: 10.1007/s11032-012-9720-0
- Kodkany, B. S., Bellad, R. M., Mahantshetti, N. S., Westcott, J. E., Krebs, N. F., Kemp, J. F., et al. (2013). Biofortification of pearl millet with iron and zinc in a randomized controlled trial increases absorption of these minerals above physiological requirements in young children. *J. Nutr.* 143, 1489–1493. doi: 10.3945/jn.113.176677
- Kramer, C. V., and Allen, S. (2015). Malnutrition in developing countries. *Paediatr Child Health.* 25, 422–427. doi: 10.1016/j.paed.2015.04.002
- Krishnan, R., and Meera, M. S. (2018). Pearl millet minerals: effect of processing on bioaccessibility. *J. Food Sci. Technol.* 55, 3362–3372. doi: 10.1007/s13197-018-3305-9
- Kulkarni, K. S., Zala, H. N., Bosamia, T. C., Shukla, Y. M., Kumar, S., Fougat, R. S., et al. (2016). *De novo* transcriptome sequencing to dissect candidate genes associated with pearl millet-downy mildew (*Sclerospora graminicola* Sacc.) interaction. *Front. Plant Sci.* 7:847. doi: 10.3389/fpls.2016.00847
- Kumar, R. R., Rai, G. K., Goswami, S., Singh, S. P., Satyavathi, C. T., and Praveen, S. (2020). *Millet the Orphan Crop for Uncertain Future*. New Delhi: Scientific International Pvt. Ltd., 194.
- Kumar, S., Hash, C., Nepolean, T., Mahendrakar, M., Satyavathi, C., Singh, G., et al. (2018). Mapping grain iron and zinc content quantitative trait loci in an inbred-derived immortal population of pearl millet. *Genes.* 9:248. doi: 10.3390/genes9050248
- Kumar, S., Hash, C. T., Thirunavukkarasu, N., Singh, G., Rajaram, V., Rathore, A., et al. (2016). Mapping quantitative trait loci controlling high iron and zinc in self and open pollinated grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 7:1636. doi: 10.3389/fpls.2016.01636
- Kumara, C. D., Bantilan, C., Rajalaxmi, A., Rai, K. N., Yadav, O. P., Gupta, S. K., et al. (2014). *Development and Diffusion of Pearl Millet Improved Cultivars in India: Impact on Growth and Yield Stability*. Patancheru: ICRIASAT.
- Lata, C. (2015). Advances in omics for enhancing abiotic stress tolerance in millets. *Proc. Indian Nat. Sci. Acad.* 81, 397–417. doi: 10.16943/ptinsa/2015/v81i2/48095
- Liang, Z., Gupta, S. K., Yeh, C. T., Zhang, Y., Ngu, D. W., Kumar, R., et al. (2018). Phenotypic data from inbred parents can improve genomic prediction in pearl millet hybrids. *G3 Genes Genom. Genet.* 8, 2513–2522. doi: 10.1534/g3.118.200242
- Liu, C. J., Witcombe, J. R., Pittaway, T. S., Nash, M., Hash, C. T., Busso, C. S., et al. (1994). An RFLP-based genetic map of pearl millet (*Pennisetum glaucum*). *Theor. Appl. Genet.* 89, 481–487. doi: 10.1007/BF00225384
- Mahendrakar, M. D., Kumar, S., Singh, R. B., Rathore, A., Potupureddi, G., Kishor, P. K., et al. (2019). Genetic variability, genotype × environment interaction and correlation analysis for grain iron and zinc contents in recombinant inbred line population of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Indian J. Genet.* 79, 545–551. doi: 10.31742/IJGPB.79.3.3
- Mahendrakar, M. D., Parveda, M., Kishor, P. K., and Srivastava, R. K. (2020). Discovery and validation of candidate genes for grain iron and zinc metabolism in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Sci. Rep.* 10, 1–16. doi: 10.1038/s41598-020-73241-7
- Manley, C., and Mayer, J. (2012). “Lipase,” in *Clinical Veterinary Advisor: Birds and Exotic Pets*, eds J. Mayer and T. M. Donnelly (W. B. Saunders; Elsevier). doi: 10.1016/B978-1-4160-3969-3.00427-3
- Manwaring, H. R., Bligh, H. F. J., and Yadav, R. (2016). The challenges and opportunities associated with biofortification of pearl millet (*Pennisetum glaucum*) with elevated levels of grain iron and zinc. *Front. Plant Sci.* 7:1944. doi: 10.3389/fpls.2016.01944
- Meena, R. P., Vishwakarma, H., Ghosh, G., Gaikwad, K., Chellapilla, T. S., Singh, M. P., et al. (2020). Novel ASR isolated from drought stress responsive SSH library in pearl millet confers multiple abiotic stress tolerance in *PgASR3* transgenic arabidopsis. *Plant Physiol. Biochem.* 156, 7–19. doi: 10.1016/j.plaphy.2020.07.031
- Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819
- Mishra, R. N., Reddy, P. S., Nair, S., Markandeya, G., Reddy, A. R., Sopory, S. K., et al. (2007). Isolation and characterization of expressed sequence tags (ESTs) from subtracted cDNA libraries of *Pennisetum glaucum* seedlings. *Plant. Mol. Biol.* 64, 713–732. doi: 10.1007/s11103-007-9193-4
- Morgan, R. N., Wilson, J. P., Hanna, W. W., and Ozias-Akin, P. (1998). Molecular markers for rust and *Pyricularia* leaf spot disease resistance in pearl millet. *Theor. Appl. Genet.* 96, 413–420.
- Moumouni, K. H., Kountche, B. A., Jean, M., Hash, C. T., Vigouroux, Y., Haussmann, B. I., et al. (2015). Construction of a genetic map for pearl millet, *Pennisetum glaucum* (L.) R. Br., using a genotyping-by-sequencing (GBS) approach. *Mol. Breed.* 35, 5–10. doi: 10.1007/s11032-015-0212-x
- Nambiar, V. S., Dhaduk, J. J., Sareen, N., Shahu, T., and Desai, R. (2011). Potential functional implications of pearl millet (*Pennisetum glaucum*) in health and disease. *J. Appl. Pharm. Sci.* 1, 62–67.
- Nantanga, K. K. M. (2006). *Lipid stabilization and partial precooking of pearl millet by thermal treatments (12-105)* (MSc (Agric.) thesis). Food Science and Technology: Department of Food Science, Faculty of Natural and Agricultural Sciences, University of 19 Pretoria. Available online at: <http://repository.up.ac.za/bitstream/handle/2263/26680/dissertation.pdf?sequence=1>
- Nayaka, S. C., Shetty, H. S., Satyavathi, C. T., Yadav, R. S., Kavi Kishor, P. B., and Nagarajud, M., et al. (2017). Draft genome sequence of *Sclerospora graminicola* the pearl millet downy mildew pathogen. *Biotechnol. Rep.* 16, 18–20. doi: 10.1016/j.btre.2017.07.006

- Neeraja, C. N., Babu, V. R., Ram, S., Hossain, F., Hariprasanna, K., Rajpurohit, B. S. et al. (2017). Biofortification in cereals: progress and prospects. *Curr Sci.* 113:6. doi: 10.18520/cs/v113/i06/1050-1057
- Obilana, A. B., and Manyasa, E. (2002). "Millets". in *Pseudocereals and Less Common Cereals: Grain Properties and Utilization Potential*, eds P. S. Belton and J. R. N. Taylor (Berlin, Heidelberg New York, NY: Springer-Verlag), 177–217.
- Pawar, V. Y., Kute, N. S., Magar, N. M., and Patil, H. T. (2018). Genotype × environment interactions for grain micronutrient contents in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *J. Pharma. Phytochem.* 7, 37–44.
- Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome*. 5, 92–102. doi: 10.3835/plantgenome2012.05.0005
- Ponnaiah, G., Gupta, S. K., Blümmel, M., Marappa, M., Pichaikannu, S., Das, R., et al. (2019). Utilization of molecular marker based genetic diversity patterns in hybrid parents to develop better forage quality multi-cut hybrids in pearl millet. *Agriculture* 9:97. doi: 10.3390/agriculture9050097
- Prakash, G., Kumar, A., Sheoran, N., Aggarwal, R., Satyavathi, C. T., Chikara, S. K., et al. (2019). First draft genome sequence of a pearl millet blast pathogen, Magnaporthe grisea strain PMg_Dl, obtained using PacBio single-molecule real-time and Illumina NextSeq 500 sequencing. *Microbiol. Resour. Announc.* 8, e01499–e01418. doi: 10.1128/MRA.01499-18
- Pujarula, V., Pusuluri, M., Bollam, S., Das, R. R., Ratnala, R., Adapala, G., et al. (2021). Genetic variation for nitrogen use efficiency traits in global diversity panel and parents of mapping populations in pearl millet. *Front. Plant Sci.* 12:625915. doi: 10.3389/fpls.2021.625915
- Punnauri, S. M., Wallace, J. G., Knoll, J. E., Hyma, K. E., Mitchell, S. E., Buckler, E. S., et al. (2016). Development of a high-density linkage map and tagging leaf spot resistance in pearl millet using genotyping-by-sequencing markers. *Plant Genome* 9, 1–13. doi: 10.3835/plantgenome2015.10.0106
- Qi, X., Lindup, S., Pittaway, T. S., Allouis, S., Gale, M. D., and Devos, K. M. (2001). Development of simple sequence repeat markers from bacterial artificial chromosome without sub-cloning. *Biotechniques* 31, 355–362.
- Qi, X., Pittaway, T. S., Lindup, S., Liu, H., Waterman, E., Padi, F. K., et al. (2004). An integrated genetic map and a new set of simple sequence repeat markers for pearl millet (*Pennisetum glaucum*). *Theor. Appl. Genet.* 109, 1485–1493. doi: 10.1007/s00122-004-1765-y
- Raheem, D., Dayoub, M., Birech, R., Nakiyemba, A. (2021). The contribution of cereal grains to food security and sustainability in Africa: potential application of UAV in Ghana, Nigeria, Uganda, and Namibia. *Urban Sci.* 5:8. doi: 10.3390/urbansci5010008
- Rai, K. N., Govindaraj, M., and Rao, A. S. (2012). Genetic enhancement of grain iron and zinc content in pearl millet. *Qual. Assur. Saf. Crops Foods*. 4, 119–125. doi: 10.1111/j.1757-837X.2012.00135.x
- Rai, K. N., Patil, H. T., Yadav, O. P., Govindaraj, M., Khairwal, I. S., Cherian, B., et al. (2014). Dhanashakti: a high-iron pearl millet variety. *Indian Farm.* 64, 32–34. Available online at: <http://oar.icrisat.org/id/eprint/8602>
- Rai, K. N., Yadav, O. P., Govindaraj, M., Pfeiffer, W. H., Yadav, H. P., Rajpurohit, B. S., et al. (2016). Grain iron and zinc densities in released and commercial cultivars of pearl millet (*Pennisetum glaucum*). *Indian J. Agri. Sci.* 86, 11–16. Available online at: https://www.researchgate.net/profile/Mahalingam_Govindaraj/publication/299597833_Grain_iron_and_zinc_densities_in_released_and_commercial_cultivars_of_pearl_millet_Pennisetum_glaucum/links/57020cca08aee995d8e0f3.pdf
- Rai, K. N., Yadav, O. P., Rajpurohit, B. S., Patil, H. T., Govindaraj, M., Khairwal, I. S., et al. (2013). Breeding pearl millet cultivars for high iron density with zinc density as an associated trait. *J. SAT Agric. Res.* 11, 1–7. Available online at: <http://oar.icrisat.org/id/eprint/7291>
- Rajaram, V., Nepolean, T., Senthilvel, S., Varshney, R. K., Vadez, V., Srivastava, R. K., et al. (2013). Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. *BMC Genomics* 14:159. doi: 10.1186/1471-2164-14-159
- Ramu, P., Deshpande, S. P., Hash, C. T., Shah, T., Upadhyaya, H. D., Riera-Lizarazu, O., et al. (2013). Population genomic and genome-wide association studies of agroclimatic traits in sorghum. *Proc. Natl. Acad. Sci. USA.* 110, 453–458. doi: 10.1073/pnas.1215985110
- Ramya, A. R., Ahamed, M. L., Satyavathi, C. T., Rathore, A., Katiyar, P., Raj, A. G., et al. (2018). Towards defining heterotic gene pools in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 8:1934. doi: 10.3389/fpls.2017.01934
- Rani, S., Singh, R., Sehrawat, R., Kaur, B. P., and Upadhyay, A. (2018). Pearl millet processing: a review. *Nutr. Food Sci.* 48, 30–44. doi: 10.1108/NFS-04-2017-0070
- Rao, P. P., Birthal, P. S., Reddy, B. V., Rai, K. N., and Ramesh, S. (2006). Diagnostics of sorghum and pearl millet grains-based nutrition in India. *Int. Sorghum Millets News Lett.* 47, 93–96.
- Reddy, P. S., Reddy, G. M., Pandey, P., Chandrasekhar, K., and Reddy, M. K. (2012). Cloning and molecular characterization of a gene encoding late embryogenesis abundant protein from *Pennisetum glaucum*: protection against abiotic stresses. *Mol. Biol. Rep.* 39, 7163–7174. doi: 10.1007/s11033-012-1548-5
- Rooney, L. W. (1978). Sorghum and pearl millet lipids. *Cereal Chem.* 55, 584–590.
- Saidou, A. A., Cloutault, J., Couderc, M., Mariac, C., Devos, K. M., and Thuillet, A. C. (2014). Association mapping, patterns of linkage disequilibrium and selection in the vicinity of the PHYTOCHROME C gene in pearl millet. *Theor. Appl. Genet.* 127, 19–32. doi: 10.1007/s00122-013-2197-3
- Saidou, A. A., Mariac, C., Luong, V., Pham, J. L., Bezançon, G., and Vigouroux, Y. (2009). Association studies identify natural variation at PHYC linked to flowering time and morphological variation in pearl millet. *Genetics* 182, 899–910. doi: 10.1534/genetics.109.102756
- Sankar, S. M., Satyavathi, C. T., Barthakur, S., Singh, S. P., Bharadwaj, C., and Soumya, S. L. (2021). Differential modulation of heat inducible genes across diverse genotypes and molecular cloning of a sHSP from Pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant Sci.* 12:659893. doi: 10.3389/fpls.2021.659893
- Satyavathi, C. T. (2017). *Proceedings of 52nd Annual Group Meet, ICAR-All India Coordinated Research Program on Pearl millet, Jodhpur held at Punjab Agricultural University Ludhiana (Punjab) during April 28–30, 2017.*
- Satyavathi, C. T., Khandelwal, V., Supriya, A., Beniwal, B. R., Sushila, B., Mahesh, C. K., et al. (2020). *Pearl Millet-Hybrids and Varieties- 2020.* Jodhpur: ICAR-All India Coordinated Research Project on Pearl Millet.
- Satyavathi, C. T., Sankar, S. M., Singh, S. P., Bhowmick, P., Bhat, J., Singh, O., et al. (2015). Stability analysis of grain iron and zinc content in pearl millet (*Pennisetum glaucum* (L.) R. Br.). *Int. J. Trop. Agric.* 33, 1387–1394.
- Satyavathi, C. T., Tiwari, S., Bharadwaj, C., Rao, A. R., Bhat, J., and Singh, S. P. (2013). Genetic diversity analysis in a novel set of restorer lines of pearl millet [*Pennisetum glaucum* (L.) R.Br] using SSR markers. *Vegetos* 26, 72–82. doi: 10.5958/j.2229-4473.26.1.011
- Sehgal, D., Rajaram, V., Armstead, I. P., Vadez, V., Yadav, Y. P., Hash, C. T., et al. (2012). Integration of gene-based markers in a pearl millet genetic map for identification of candidate genes underlying drought tolerance quantitative trait loci. *BMC Plant Biol.* 12:9. doi: 10.1186/1471-2229-12-9
- Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., Taunk, J., et al. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS ONE* 10:e0122165. doi: 10.1371/journal.pone.0122165
- Senthilvel, S., Jayashree, B., Mahalakshmi, V., Kumar, P. S., Nakka, S., Nepolean, T., et al. (2008). Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. *BMC Plant Biol.* 8:119. doi: 10.1186/1471-2229-8-119
- Senthilvel, S., Nepolean, T., Supriya, A., Rajaram, V., Kumar, S., Hash, C. T., et al. (2010). "Development of a molecular linkage map of pearl millet integrating DArT and SSR markers". in *Proc Plant Animal Genome 18 Conference* (San Diego, CA), 9–13.
- Serba, D. D., Muleta, K., Amand, P., St. Bernardo, A., Bai, G., Perumal, R., et al. (2019). Genetic diversity, population structure, and linkage disequilibrium of pearl millet. *Plant Genom.* 12:180091. doi: 10.3835/plantgenome2018.11.0091
- Serba, D. D., and Yadav, R. S. (2016). Genomic tools in pearl millet breeding for drought tolerance: status and prospects. *Front. Plant Sci.* 7:1724. doi: 10.3389/fpls.2016.01724
- Serba, D. D., Yadav, R. S., Varshney, R. K., Gupta, S. K., Mahalingam, G., Srivastava, R. K., et al. (2020). "Genomic designing of pearl millet: a resilient crop for arid and semi-arid environments", in *Genomic Designing of Climate-Smart Cereal Crops*, ed C. Kole (Springer Nature), 221–286. doi: 10.1007/978-3-319-93381-8_6
- Serraj, R., Hash, C. T., Rizvi, S. M., Sharma, A., Yadav, R. S., and Bidinger, F. R. (2005). Recent advances in marker-assisted selection for drought tolerance in pearl millet. *Plant Prod. Sci.* 8, 334–337. doi: 10.1626/pp.s.8.334
- Sharma, P. C., Singh, D., Sehgal, D., Singh, G., Hash, C. T., and Yadav, R. S. (2014). Further evidence that a terminal drought tolerance QTL of pearl millet

- is associated with reduced salt uptake. *Environ. Exp. Botany*. 102, 48–57. doi: 10.1016/j.envexpbot.2014.01.013
- Sharma, S., Sharma, R., Govindaraj, M., Mahala, R. S., Satyavathi, C. T., Srivastava, R. K., et al. (2021). Harnessing wild relatives of pearl millet for germplasm enhancement: challenges and opportunities. *Crop Sci. (TSI)*. 61, 177–200. doi: 10.1002/csc.2.20343
- Shivhare, R., Asif, M. H., and Lata, C. (2020). Comparative transcriptome analysis reveals the genes and pathways involved in terminal drought tolerance in pearl millet. *Plant Mol. Biol.* 103, 639–652. doi: 10.1007/s11103-020-01015-w
- Siddaiah, C. N., Prasanth, K. V. H., Satyanarayana, N. R., Mudili, V., Gupta, V. K., Kalagatur, N. K., et al. (2018). Chitosan nanoparticles having higher degree of acetylation induce resistance against pearl millet downy mildew through nitric oxide generation. *Sci. Rep.* 8, 1–14. doi: 10.1038/s41598-017-19016-z
- Singh, J., Reddy, P. S., Reddy, S., and Reddy, M. (2015). Molecular cloning and characterization of salt inducible dehydrin gene from the C4 plant *Pennisetum glaucum*. *Plant Gene*. 4, 55–63. doi: 10.1016/j.plgene.2015.08.002
- Singh, N., Singh, S. P., Kumar, M., Kanhiya, K., and Kumar, A. (2018). Nutricereal pearl millet: way forward. *Int. J. Curr. Microbiol. App. Sci.* 7, 578–581. doi: 10.20546/ijcmas.2018.706.066
- Singh, S. D., Wilson, J. P., Navi, S. S., Talukdar, B. S., Hess, D. E., and Reddy, K. N. (1997). Screening techniques and sources of resistance to downy mildew and rust in pearl millet. Information Bulletin No. 48. *Int. Crops Res. Instit. Semi-Arid Trop. Patancheru Andhra Pradesh India*. 36:110.
- Singh, S. P., Satyavathi, C. T., and Sankar, S. M. (2014). “Hybrid breeding in pearl millet: past and present status”, in H. S. Chawla, P. K. Shrotria, and A. S. Jeena, eds *New Paradigms in Heterosis Breeding: Conventional and Molecular Approaches* (New Delhi: Indian Agric Stat Res Inst).
- Singhal, T., Satyavathi, C. T., Kumar, A., Sankar, S. M., Singh, S. P., Bharadwaj, C., et al. (2018). Genotype \times environment interaction and genetic association of grain iron and zinc content with other agronomic traits in RIL population of pearl millet. *Crop Pasture Sci.* 69, 1092–1102. doi: 10.1071/CP18306
- Singhal, T., Satyavathi, C. T., Singh, S. P., Kumar, A., Sankar, S. M., Bharadwaj, C., et al. (2021). Multi environment quantitative trait loci mapping for grain iron and zinc content using biparental recombinant inbred line population in pearl millet. *Front. Plant Sci.* 12:659789. doi: 10.3389/fpls.2021.659789
- Srivastava, R. K., Bollam, S., Pujarula, V., Pusuluri, M., Singh, R. B., Potupureddi, G., et al. (2020b). Exploitation of heterosis in pearl millet: a review. *Plants*. 9:807. doi: 10.3390/plants9070807
- Srivastava, R. K., Singh, R. B., Lakshmi, V., Srikanth P., Pusuluri B., Satyavathi M., et al. (2020a). Genome wide association studies and genomic selection in Pearl millet: Advances and Prospects. *Front. Genet.* 10:1389. doi: 10.3389/fgene.2019.01389
- Stein, A. J., Qaim, M., Meenakshi, J. V., Nestel, P., Sachdev, H. P. S., and Bhutta, Z. A. (2007). “Biofortification, an agricultural micronutrient intervention: its potential impact and cost-effectiveness”. in: *Consequences and Control of Micronutrient Deficiencies: Science, Policy, and Programs, April 16-18 (Istanbul)*.
- Sun, M., Huang, D., Zhang, A., Khan, I., Yan, H., Wang, X., et al. (2020). Transcriptome analysis of heat stress and drought stress in pearl millet based on Pacbio full-length transcriptome sequencing. *BMC Plant Biol.* 20, 1–15. doi: 10.1186/s12870-020-02530-0
- Supriya, A., Senthilvel, S., Nepolean, T., Eshwar, K., Rajaram, V., Shaw, R., et al. (2011). Development of a molecular linkage map of pearl millet integrating DArT and SSR markers. *Theor. App. Genet.* 123, 239–250. doi: 10.1007/s00122-011-1580-1
- Taunk, J., Sehgal, D., Yadav, N. R., Howarth, C., Yadav, R. C., and Yadav, R. S. (2018). Mapping of easy to screen SSR markers for selection of RFLP markers-bracketed downy mildew resistance QTLs in pearl millet. *Eur. J. Plant Pathol.* 151, 401–411. doi: 10.1007/s10658-017-1381-8
- Thakur, R. P., and King, S. B. (1988). Registration of six smut resistant germplasms of pearl millet. *Crop Sci.* 28, 382–383. doi: 10.2135/cropsci1988.0011183X002800020058x
- Thakur, R. P., Rai, K. N., Khairwal, I. S., and Mahala, R. S. (2008). Strategy for downy mildew resistance breeding in pearl millet in India. *J SAT Agric Res.* 6, 1–11.
- Tharanya, M., Kholova, J., Sivasakthi, K., Seghal, D., Hash, C. T., Raj, B., et al. (2018). Quantitative trait loci (QTLs) for water use and crop production traits co-locate with major QTL for tolerance to water deficit in a fine mapping population of pearl millet (*Pennisetum glaucum* L. R.Br.). *Theor. Appl. Genet.* 131:1509–29. doi: 10.1007/s00122-018-3094-6
- Uppal, R. K., Wani, S. P., Garg, K. K., and Alagaraswamy, G. (2015). Balanced nutrition increases yield of pearl millet under drought. *Field Crop Res.* 177, 86–97. doi: 10.1016/j.fcr.2015.03.006
- Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35:969. doi: 10.1038/nbt.3943
- Velu, G., Rai, K. N., Muralidharan, V., Kulkarni, V. N., Longvah, T., and Raveendran, T. S. (2007). Prospects of breeding biofortified pearl millet with high grain iron and zinc content. *Plant Breed.* 126, 182–185. doi: 10.1111/j.1439-0523.2007.01322.x
- Verma, D., Singla-Pareek, S. L., Rajagopal, D., Reddy, M. K., and Sopory, S. K. (2007). Functional validation of a novel isoform of Na⁺/H⁺ antiporter from *Pennisetum glaucum* for enhancing salinity tolerance in rice. *J. Biosci.* 32, 621–628. doi: 10.1007/s12038-007-0061-9
- Weckwerth, W., Ghatak, A., Bellaire, A., Chaturvedi, P., and Varshney, R. K. (2020). PANOMICS meets germplasm. *Plant Biotechnol. J.* 18, 1507–1525. doi: 10.1111/pbi.13372
- Yadav, O. P., and Rai, K. N. (2013). Genetic improvement of pearl millet in India. *Agric. Res.* 2, 275–292. doi: 10.1007/s40003-013-0089-z
- Yadav, O. P., Gupta, S. K., Govindara, J. M., Sharma, R., Varshney, R. K., Srivastava, R. K., et al. (2021). Genetic gains in pearl millet in India: insights into historic breeding strategies and future perspective. *Front. Plant Sci.* 12:645038. doi: 10.3389/fpls.2021.645038
- Yadav, O. P., Singh, D. V., Dhillon, B. S., and Mohapatra, T. (2019). India's evergreen revolution in cereals. *Curr. Sci.* 116, 1805–1808. doi: 10.18520/cs/v116/i11/1805-1808
- Yadav, R. S., Hash, C. T., Bidinger, F. R., Cavan, G. P., and Howarth, C. J. (2002). Quantitative trait loci associated with traits determining grain and stover yield in pearl millet under terminal drought stress conditions. *Theor. Appl. Genet.* 104, 67–83. doi: 10.1007/s001220200008
- Yadav, R. S., Hash, C. T., Bidinger, F. R., Devos, K. M., and Howarth, C. J. (2004). Genomic regions associated with grain yield and aspects of post flowering drought tolerance in pearl millet across stress environments and tester background. *Euphytica*. 136, 265–277. doi: 10.1023/B:EUPH.0000032711.34599.3a
- Yadav, R. S., Hash, C. T., Bidinger, F. R., and Howarth, C. J. (1999). “QTL analysis and marker-assisted breeding of traits associated with drought tolerance in pearl millet”. in *Genetic Improvement of Rice for Water-Limited Environments*, eds O. Ito, J. O'Toole, and B. Hardy (Manila: International Rice Research Institute), 221–233.
- Yadav, R. S., Sehgal, D., and Vadez, V. (2011). Using genetic mapping and genomics approaches in understanding and improving drought tolerance in pearl millet. *J. Exp. Bot.* 62, 397–408. doi: 10.1093/jxb/erq265

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Satyavathi, Ambawat, Khandelwal and Srivastava. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Identification of High-Yielding Iron-Biofortified Open-Pollinated Varieties of Pearl Millet in West Africa

Prakash I. Gangashetty^{1,2*}, Mohammed Riyazaddin², Moussa Daouda Sanogo³, Drabo Inoussa⁴, Kassari Ango Issoufou⁵, Peter A. Asungre⁶, Ousmane Sy⁷, Mahalingam Govindaraj^{1*} and Angarawai Ijantiku Ignatius⁸

¹ International Crops Research Institute for the Semi-Arid Tropics, (ICRISAT), Patancheru, India, ² International Crops Research Institute for the Semi-Arid Tropics, (ICRISAT), Niamey, Niger, ³ Institut d'Economie Rurale, Cinzana, Mali, ⁴ Institut de l'Environnement et du Recherches Agricoles de Burkina Faso (INERA), Ouagadougou, Burkina Faso, ⁵ Institut National de la Recherche Agronomique du Niger (INRAN), Maradi, Niger, ⁶ CSIR-Savanna Agriculture Research Institute, Bawku, Ghana, ⁷ Senegalese Institute of Agricultural Research (ISRA), Bambey, Senegal, ⁸ International Crops Research Institute for the Semi-Arid Tropics, (ICRISAT), Kano, Nigeria

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Karthikeyan Adhimoolam,
Jeju National University, South Korea
Om Parkash Yadav,
Central Arid Zone Research Institute
(ICAR), India
Sushil Kumar,
Anand Agricultural University, India

*Correspondence:

Prakash I. Gangashetty
p.gangashetty@cgiar.org
Mahalingam Govindaraj
m.govindaraj@cgiar.org

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 31 March 2021

Accepted: 26 July 2021

Published: 22 September 2021

Citation:

Gangashetty PI, Riyazaddin M, Sanogo MD, Inoussa D, Issoufou KA, Asungre PA, Sy O, Govindaraj M and Ignatius AI (2021) Identification of High-Yielding Iron-Biofortified Open-Pollinated Varieties of Pearl Millet in West Africa. *Front. Plant Sci.* 12:688937. doi: 10.3389/fpls.2021.688937

Pearl millet is a predominant food and fodder crop in West Africa. This study was carried out to test the newly developed open-pollinated varieties (OPVs) for field performance and stability for grain yield, grain iron (Fe), and grain zinc (Zn) contents across 10 locations in West Africa (i.e., Niger, Nigeria, Mali, Burkina Faso, Senegal, and Ghana). The test material consisted of 30 OPVs, of which 8 are Fe/Zn biofortified. The experiment was conducted in a randomized complete block design in three replications. ANOVA revealed highly significant variability for grain yield and micronutrient traits. The presence of genotype \times environment ($G \times E$) indicated that the expressions of traits are significantly influenced by both genetic and $G \times E$ factors, for grain Fe and Zn contents. Days to 50% flowering and plant height showed less $G \times E$, suggesting these traits are largely under genetic control. The genotypes CHAKTI (46 days), ICTP 8203 (46 days), ICMV 177002 (50 days), ICMV 177003 (48 days), and Moro (53 days) had exhibited early flowering across locations leading to early physiological maturity. CHAKTI (1.42 t/ha yield; 62.24 mg/kg of grain Fe, 47.29 mg/kg of grain Zn) and ICMP 177002 (1.19 t/ha yield, 62.62 mg/kg of grain Fe, 46.62 mg/kg of grain Zn) have performed well for grain yield and also for micronutrients, across locations, compared with the check. Additive Main Effect and Multiplicative Interaction (AMMI) ANOVA revealed the highly significant genotypic differences, the mean sum of squares of environment, and its interaction with the genotypes. Based on the AMMI stability value (ASV), the most stable genotype is SOSAT-C88 (ASV = 0.04) for grain yield and resistance to downy mildew; mean grain yield and stability rankings (YSI) revealed that the genotypes CHAKTI, SOSAT-C88, and ICMV IS 99001 were high yielding and expressed stability across regions. The strong correlation ($r = 0.98^{**}$) of grain Fe and Zn contents that merits Fe-based selection is highly rewarding. CHAKTI outperformed over other genotypes for grain yield (71%

higher), especially with early maturing varieties in West Africa, such as GB 8735, LCIC 9702, and Jirani, and for grain Fe (16.11% higher) and Zn (7% higher) contents across locations, and made a candidate of high-iron variety to be promoted for combating the micronutrient malnutrition in West and Central Africa (WCA).

Keywords: pearl millet, micronutrient malnutrition (MNM), CHAKTI, AMMI analysis, grain Fe and Zn

INTRODUCTION

Pearl millet (*Pennisetum glaucum* L.) is an important food and fodder crop for people living in the semi-arid regions of Asia and Africa. The people living in the semi-arid regions mostly consume cereal-based foods and do not have access to diverse food, hence, leading to serious malnutrition in children and anemia in women (Underwood, 2000). Micronutrient malnutrition is a major concern in the developing world, where more than 2 billion people are malnourished (World Health Organization, 2006; Tulchinsky, 2010; Kramer and Allen, 2015; Webb et al., 2018). Malnutrition in Africa is a rising concern where children are affected by stunted growth and low IQ. Young children and pregnant women are most vulnerable to micronutrient deficiencies due to their rapid growth and development (Black et al., 2013; Saltzman et al., 2017). At least half of the children worldwide younger than 5 years of age suffer from vitamin and mineral deficiencies, while almost 2 in 3 children between 6 months and 2 years of age are not fed on food that supports their healthy development (UNICEF, 2019). Although only required in small amounts, micronutrients are not produced in the body and must be derived from the diet. It is severe in regions where people depend on staple crops, which are low in micronutrients, to meet their energy requirements. Hence, improving the micronutrient concentration of locally adapted staple food crops will be a sustainable solution to address this malnutrition. Efforts are being made to provide fortified foods to these vulnerable groups, which is a cost-effective strategy, but it is difficult to ensure the availability of these fortified foods to people living in remote areas. Food diversification is another tool to combat micronutrient malnutrition, but the high cost involved and limited food availability make it unsustainable.

Biofortification, the process of increasing the concentration of micronutrients in the staple crops to have a measurable impact on human nutrition after consumption, is therefore important for combating micronutrient malnutrition. Biofortification is becoming more of the breeding process by which the nutrient density is increased through conventional plant breeding and modern biotechnology without altering the preferred traits of the farmers. (1) Iron biofortification of beans, cowpea, and pearl millet, (2) zinc biofortification of maize, rice, and wheat, and provitamin A, and (3) carotenoid biofortification of cassava, maize, rice, and sweet potato are currently underway and at different stages of development (Saltzman et al., 2013). Biofortification has been regarded as a novel and sustainable vehicle in the reduction of micronutrient malnutrition (Bouis et al., 2011).

The HarvestPlus program of the CGIAR is supporting global biofortification efforts, through which the International Crops

Research Institute for the Semi-Arid Tropics (ICRISAT) released the ever-first high grain Fe pearl millet variety “Dhanashakti” in 2014 by exploiting the intrapopulation variability (recurrent selection) within ICTP 8203 variety, an early open-pollinated variety (OPV) popularly grown in India (Rai et al., 1990; Govindaraj et al., 2019). Since then, pearl millet hybrids rich in kernel iron and zinc concentrations have been bred and released in India through conventional (i.e., heterosis) breeding methods (Govindaraj et al., 2018, 2019).

Many studies have shown the positive impact of the consumption of biofortified crops, and the results of these studies were also similar to iron fortification and supplementation strategies to fight against hidden hunger (Kodkany et al., 2013; Saltzman et al., 2013; Haas, 2014; Finkelstein et al., 2015). The use of biofortified pearl millet varieties could have a positive effect on Fe and Zn intakes in the long term. The studies conducted by Kodkany et al. (2013) revealed that the quantities of both iron and zinc were absorbed when iron- and zinc-biofortified pearl millet was fed to children aged 2 years as the major staple food and was more than adequate to meet the physiological requirements for these micronutrients.

The regions in West and Central Africa (WCA), where pearl millet is significantly grown and consumed, have no biofortified variety at present. The ICRISAT center in Niamey, Niger with the help of the HarvestPlus program employs in developing high-yielding biofortified pearl millet varieties and hybrids. WCA regions significantly hold the undernutrition populations, and it is very appropriate to breed, test, and identify the Fe/Zn-rich pearl millet varieties besides better yield and resistance to biotic stress across the region of WCA. Most of the pearl millet varieties existing with the farming community have lower levels of grain Fe (<30.0 mg/kg) and Zn (<30.0 mg/kg) contents. Hence, this study was conducted to identify the pearl millet varieties with high grain Fe and Zn contents, their stable performance across the region, and the influence of environmental factors on the inheritance of these traits.

MATERIALS AND METHODS

Experimental Material and Design

The experimental material consisted of 30 pearl millet varieties, of which 8 varieties had high grain Fe and Zn contents, 20 with moderate levels of grain Fe and Zn, and 2 local checks GB 8735 and Jirani. GB 8735 is the highest acreage cultivated in Niger, Togo, and adjoining countries with its early maturity and stability. This variety was also noticed with higher Fe/Zn in the regions. Using this variety for both early maturity and micronutrient is highly appropriate in this

study. The trials were conducted in 10 locations of WCA representing the Sahelian agroecology with an annual rainfall of 400 mm. The test locations were Niger (Sadore, Konni, Maradi, and Magaria), Nigeria (Kano), Mali (Cinzana), Burkina Faso (Gampella), Ghana (Bawku), and Senegal (Bambey and Nioro) in WCA. The material was sown in randomized complete block design in three replications in the rainy season of 2018. A basal dose of fertilizer was applied to the field at the rate of 100 kg/ha. The genotypes were sown in two-row plots with a row length of 3.0 m long and a row-to-row distance of 0.80 and 0.20 m in between the plants. Thinning of the test entries to one plant per hill was carried out at 15–21 days after seedling emergence (DAE). Hand weeding of the test plots was carried out when necessary. Microdosing with urea was carried out at 30 DAE.

Data on agronomic and morphological traits were recorded in the test plots. Data on days to 50% flowering, plant height (cm), number of downy mildew plants, panicle length (cm), circumference (cm), and grain yield (t/ha) were recorded. Downy mildew incidence were calculated based on the percentage incidence using the number of plants affected to the total number of plants in the plot multiplied by 100. Grain samples for Fe and Zn analysis were collected before harvesting on the standing crop eliminating the plants grounded and contaminated with soil. Proper care was taken in cleaning the grain samples for analysis to avoid contamination from metals and dust.

Micronutrient Estimation

Pearl millet grain micronutrient estimation was carried out using the energy dispersive X-ray fluorescence (EDXRF) (Paltridge et al., 2012a,b; Guild and Stangoulis, 2016) method. The EDXRF analysis was carried out using Oxford X-Supreme 8000 (Oxford Instruments plc, United Kingdom) with a 10-sample carousel. Analyses were conducted in supplied sample cups prepared as reported previously (Paltridge et al., 2012a,b; Guild and Stangoulis, 2016), with a 4- μ m Poly-4 XRF sample film used to seal one end of the cup. Sample cups were cleaned and prepared before each analysis to minimize cross-contamination between samples. Samples of >5 g were used for all analyses to ensure that samples were “infinitely thick” in terms of EDXRF analysis (Paltridge et al., 2012b). Calibrations for Fe and Zn in pearl millet were achieved using 20 standards based on the data of inductively coupled plasma optical emission spectrophotometry (ICP-OES).

Statistical Analysis

The replicated data from each location are subjected to perform ANOVA and Additive Main Effects and Multiplicative Interaction (AMMI) analysis using GenStat software version 18.0 (GenStat, 2015). The significant differences between the genotypes were tested by using the *F*-test, and the genotypic means were compared by the least significant difference at $p \leq 0.05$. AMMI analysis permits the estimation of the interaction effect of a genotype in each environment, and it helps to identify genotypes best suited for specific environmental conditions. The trait phenotypic correlation coefficient was analyzed using Excel.

Heritability estimates were calculated using $H^2 = V_g/V_p$, where H^2 is the heritability estimate; V_g , the variation in genotype; and V_p , the variation in phenotype.

RESULTS

The ANOVA has revealed a highly significant ($P \leq 0.01$) mean sum of squares of entries, locations (hereafter referred to as environment), and genotypes \times location [hereafter referred to as genotype \times environment ($G \times E$)] for all the studied traits (Table 1). The mean sum of squares of genotypes was higher than the mean sum of squares of environments. The $G \times E$ influenced the trait determination in these 10-environment testing; however, it was very much lower in variance magnitude for all the traits. For instance, flowering, grain yield, and grain Fe and Zn contents were largely controlled by genotypic variance. Days to 50% flowering showed 29 times higher mean square magnitude than the $G \times E$ source of variation. The heritability for agronomic and morphological traits revealed presence of high heritability for all the studied traits. The estimation of genetic advance percentage mean showed high genetic advance for downy mildew resistance (106.15%), panicle length (66.50%), grain yield (49.32%), and grain Fe content (39.01%).

The mean performance of the pearl millet genotypes indicated significant *F* probabilities ($P \leq 0.01$) for all the studied traits (Table 2). The genotypes CHAKTI, Faringuero, IBMV 8402, ICMV 167005, ICMV 167006, ICMV 177001, ICMV 177002, ICMV 177003, ICMV 177004, ICMV IS 85327, ICMV IS 99001, ICTP 8203, Moro, SARIA-1, SOSAT-C88, Souna 3, and WAAP-NAARA had shown lower (<10%) downy mildew damage percentage when compared with the susceptible check, Sadore local (17.69 %).

The genotypes CHAKTI, ICMV 221, ICTP 8203, and Jirani had attained days to 50% flowering within 45 days, which is very early (Table 2). Other 11 genotypes also exhibited lower days to 50% flowering when compared with the overall mean (55.00 days) and with the local check (66.60 days). The mean plant height of 202.00 cm is recorded across locations. Fourteen genotypes exhibited greater plant height when compared with the mean plant height. Most of the genotypes with the lower days to 50% flowering exhibited moderate plant height.

The test genotypes exhibited a mean panicle length of 32.00 cm across locations (Table 2). The genotypes Gamoji (41.10 cm, 9.91 cm; panicle length and panicle circumference, respectively) and ICMV 177004 (32.97, 9.12) have shown greater panicle length and circumference across locations. The genotype ICMV IS 94222 had exhibited the highest panicle length of 60.01 cm, and AKAD-KOM has the least panicle length of 17.56 cm.

The mean performance of genotypes for grain yield across locations is given in Table 2. The overall mean for grain yield is 0.96 t/ha. The genotype CHAKTI (1.42 t/ha) has performed well for grain yield when compared with the other genotypes. The genotypes Gamoji, IBMV 8402, ICMV 177002, ICMV 167005, ICMV 167006, ICMV 167001, ICMV IS 85327, ICMV IS 94222, ICMV IS 99001, SOSAT-C88, and Souna 3 yielded high across locations when compared with the check Sadore local (0.73 t/ha).

TABLE 1 | Mean sum of squares of pearl millet genotypes across locations in West Africa.

Sources of variation	df	Downy mildew damage	Days to 50% flowering	Plant height	Panicle length	Panicle circumference	Grain yield	Grain Fe content	Grain Zn content
Replication	2	363.00	11.81	616.50	61.34	4.58	0.02	12.72	36.54
Environment	9	9924.10**	660.53**	7793.60**	334.48**	14.65**	1.62**	197.27**	361.60**
Genotype	29	1369.10**	1628.99**	26451.30**	3316.05**	47.92**	1.75**	2125.37**	689.32**
Environment × genotype	260	296.50**	56.04**	757.00**	80.04**	1.40**	0.12**	27.43**	28.12**
Residual	574	102.90	9.61	233.90	15.36	1.04	0.02	8.95	12.39
Total	892								
Heritability (H^2)		80.21	96.62	97.18	97.62	97.09	93.39	98.72	95.98
Genetic advance as % mean		106.15	26.92	29.62	66.50	28.16	49.32	39.01	25.81

**Mean sum of squares significant at $P = 0.01$.

TABLE 2 | Mean performance of the pearl millet (*Pennisetum glaucum* L.) genotypes across locations in West Africa.

Genotype	Downy mildew damage%	Days to 50% flowering	Plant height (cm)	Panicle length (cm)	Panicle circumference (cm)	Grain yield (t/ha)	Grain Fe content (mg/kg)	Grain Zn content (mg/kg)
AFRIBEH-NAARA	12.09	47.43	190.00	24.03	9.30	0.79	45.04	38.60
AKAD-KOM	15.92	48.03	177.40	17.56	10.45	0.83	52.66	42.39
CHAKTI	1.97	45.57	168.50	23.78	9.59	1.42	62.24	47.29
Faringuero	9.57	58.87	208.20	35.58	8.99	0.73	34.55	32.31
Gamoji	11.13	62.57	231.40	41.10	9.91	1.19	37.29	32.82
GB 8735	17.22	49.43	170.30	23.39	10.22	0.83	53.60	44.02
IBMV 8402	8.29	61.97	210.30	39.02	7.83	0.99	38.21	33.86
ICMP 177001	23.53	48.63	171.30	23.71	10.53	0.85	57.67	43.77
ICMP 177002	26.25	48.37	173.20	25.24	10.03	1.19	62.62	46.62
ICMV 167005	3.27	65.63	240.60	44.28	8.08	1.31	38.12	35.70
ICMV 167006	6.88	59.43	231.50	32.10	6.95	1.13	42.41	35.51
ICMV 177001	4.14	65.73	256.70	29.61	7.19	0.61	37.51	32.41
ICMV 177002	3.74	50.07	195.30	26.84	9.34	0.83	43.72	36.89
ICMV 177003	9.92	48.40	168.40	23.83	10.29	0.83	52.17	40.79
ICMV 177004	9.82	56.23	179.00	32.97	9.12	0.80	40.70	36.01
ICMV 221	11.56	44.57	163.70	22.93	10.12	0.79	51.36	42.63
ICMV 167001	23.01	49.07	176.60	29.46	10.96	1.33	50.52	40.23
ICMV IS 85327	10.94	59.87	234.10	51.06	7.93	0.97	37.05	33.00
ICMV IS 94222	11.88	60.57	241.70	60.01	7.80	1.28	35.60	32.86
ICMV IS 99001	5.63	58.93	218.00	51.12	7.87	1.32	33.34	30.24
ICTP 8203	9.95	45.47	168.80	24.79	10.29	0.75	55.07	43.19
Jirani	12.66	45.73	172.20	22.64	8.35	0.84	47.18	39.93
KAANATI	20.12	54.43	190.30	28.61	8.60	0.69	43.04	35.81
LCIC 9702	12.02	49.10	184.40	30.09	8.66	0.87	43.59	35.57
Moro	8.93	53.00	209.80	28.38	7.17	1.06	45.33	37.43
Sadore Local	17.69	66.60	235.50	46.46	7.29	0.73	33.71	29.84
SARIA-1	4.24	69.90	260.80	30.58	7.81	0.61	43.61	37.30
SOSAT-C88	1.19	58.27	221.60	28.19	11.11	1.38	36.64	33.16
Souna 3	2.26	60.30	221.40	47.95	7.59	0.98	38.39	34.91
WAAPP-NAARA	2.34	56.00	192.90	26.03	10.09	0.88	37.50	33.56
Mean	11.00**	55.00**	202.00**	32.00**	9.00**	0.96**	44.00**	37.00**
Vr	13.31	169.49	113.07	215.90	45.99	88.67	237.47	55.65
SE ±	1.90	0.60	2.80	0.70	0.20	0.03	0.50	0.60
LSD	5.10	1.60	7.80	2.00	0.50	0.07	1.50	1.80
CV (%)	95.60	5.60	7.60	12.10	11.40	14.60	6.70	9.40

**F probabilities significant at $P = 0.01$; Vr, variance ratio; SE, standard error; LSD, least significant difference; CV, coefficient of variation.

TABLE 3 | Analysis of variance for AMMI model of pearl millet genotypes across locations in West Africa.

Source	df	Downy mildew damage	Days to 50% flowering	Plant height	Panicle length	Panicle circumference	Grain yield	Grain Fe content	Grain Zn content
Treatments	299	689.00**	226.80**	3461.00**	401.60**	6.31**	0.33**	236.00**	102.30**
Genotypes (G)	29	1369.00**	1629.00**	26451.00**	3316.10**	47.92**	1.75**	2125.40**	689.30**
Environments (E)	9	9924.00**	660.50**	7794.00**	334.50**	14.65**	1.62**	197.30**	361.60**
Block	20	305.00**	10.80 ^{NS}	394.00*	19.90 ^{NS}	1.88*	0.02 ^{NS}	41.00**	28.10**
G × E	260	296.00**	56.00**	757.00**	80.00**	1.40**	0.12**	27.40**	28.10**
IPCA 1	37	1067.00**	248.50**	2345.00**	270.60**	3.18**	0.35**	53.30**	61.50**
IPCA 2	35	496.00**	49.90**	1099.00**	164.80**	3.14**	0.12**	45.80**	40.70**
Residuals	188	108.00	19.50	383.00**	27.00	0.73	0.08	19.00	19.30
Error	574	103.00	9.60	234.00	15.40	1.04	0.02	9.00	12.40

*, **F probabilities significant at $P = 0.05$ and 0.01 , respectively; NS, non-significant probabilities; IPCA 1, interaction principal component axis one; IPCA 2, interaction principal component axis two.

The genotypes AFRIBEH-NAARA, AKAD-KOM, GB 8735, ICMP 177001, ICMV 177003, ICMV 221, ICTP 8203, and Jirani exhibited high grain Fe and Zn contents across locations when compared with the check Sadore local (33.71 mg/kg Fe and 29.84 mg/kg Zn), expressing lower than the overall mean (Table 2). The genotypes CHAKTI and ICMP 177002 have the highest grain Fe content of >62.0 mg/kg and grain Zn content of >45.0 mg/kg.

The AMMI ANOVA carried out to partition the total variances into its components revealed existence of highly significant ($P \leq 0.01$) genotypic differences among all the traits (Table 3). The mean sum of squares of the environment and its interaction with the genotypes were also significant at $P \leq 0.01$. The $G \times E$ interaction was further partitioned into IPCA 1 and IPCA 2, which were significant at $P \leq 0.01$ for all the studied traits. The mean sum of squares of $G \times E$ interactions was lower than the mean sum of squares of the genotypes.

In the AMMI biplot, the IPCA 1 (PC1) and IPCA 2 (PC2) scores of genotype and environments were plotted against each other for grain yield (Figure 1) and grain Fe content (Figure 2). The IPCA 1 component accounted for 40.26 and 27.54% of $G \times E$ interaction, while IPCA 2 accounted for 12.75 and 22.40%, respectively, for grain yield and grain Fe content. In the biplot for grain yield, the genotypes ICMP 177002, IBMV 8402, Souna 3, ICMV IS 85327, and Moro are scattered away from the origin and away from the environmental spikes. Most of the genotypes are close to the origin in the biplot for grain Fe content. The genotypes AFRIBEH-NAARA, ICMV 177001, ICMV 177002, and SOSAT-C88 are scattered away from the origin and also away from the environment spikes.

The AMMI stability value (ASV) ranked the genotypes based on the least score (Table 4). Low scores represent the most stable genotypes across locations. Based on the ASV, the most stable genotype is SOSAT-C88 (ASV = 0.04) followed by LCIC 9702 (ASV = 0.13) for grain yield, ICMV 221 (ASV = 0.05) followed by ICMV IS 99001 (ASV = 0.23) for grain Fe content, and WAAP-NAARA (ASV = 0.28) followed by GB 8735 (ASV = 0.50) for grain Zn content. The sum of mean grain yield and stability rankings (YSI) revealed that the genotypes SOSAT-C88, ICMV IS 99001, and CHAKTI were high yielding and expressed stability

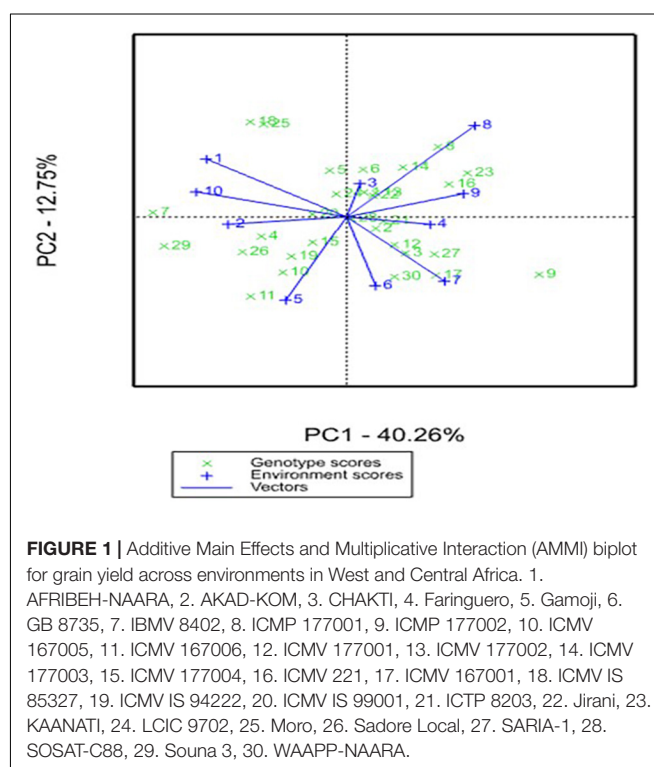
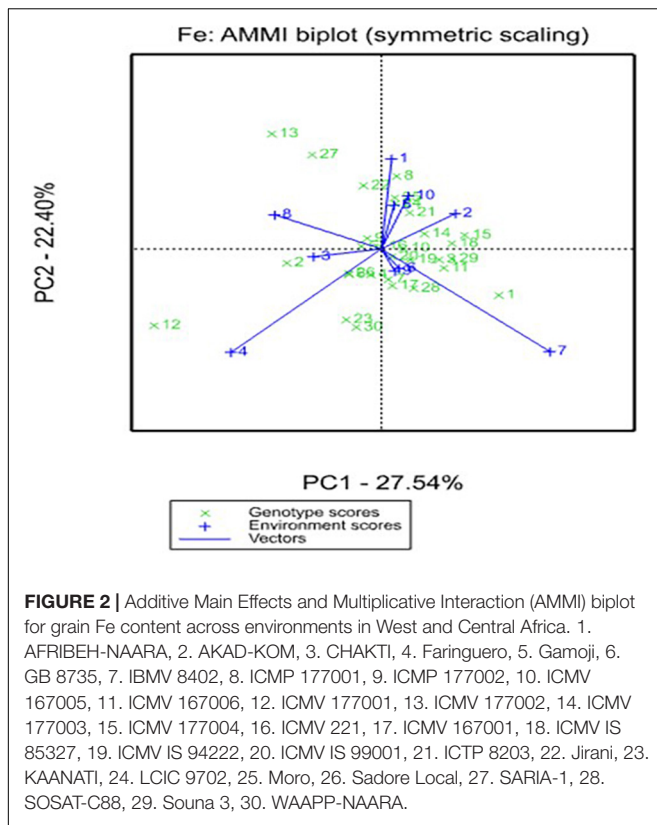


FIGURE 1 | Additive Main Effects and Multiplicative Interaction (AMMI) biplot for grain yield across environments in West and Central Africa. 1. AFRIBEH-NAARA, 2. AKAD-KOM, 3. CHAKTI, 4. Faringuero, 5. Gamoji, 6. GB 8735, 7. IBMV 8402, 8. ICMP 177001, 9. ICMP 177002, 10. ICMV 167005, 11. ICMV 167006, 12. ICMV 177001, 13. ICMV 177002, 14. ICMV 177003, 15. ICMV 177004, 16. ICMV 221, 17. ICMV 167001, 18. ICMV IS 85327, 19. ICMV IS 94222, 20. ICMV IS 99001, 21. ICTP 8203, 22. Jirani, 23. KANATI, 24. LCIC 9702, 25. Moro, 26. Sadore Local, 27. SARIA-1, 28. SOSAT-C88, 29. Souna 3, 30. WAAPP-NAARA.

across regions. The genotypes ICMV 167001 and ICMV 167005 were high yielding but were unstable, which has a high YSI rank.

Additive main effects and multiplicative interaction analysis identified four high-yielding genotypes in each of the environments (Table 5). CHAKTI performed well and yielded high at Bambey (Senegal), Niore (Senegal), Kano (Nigeria), Konni (Niger), Bawku (Ghana), and Gampella (Burkina Faso). The genotype ICMV 167005 performed well at Magaria, Sadore in Niger, and Cinzana location of Mali.

The phenotypic correlation patterns of micronutrient traits of pearl millet indicated the presence of a highly significant and positive correlation between grain Fe and Zn contents ($r = 0.98$; $P < 0.01$) (Table 6). The grain Fe and Zn contents



exhibited significant and negative correlations with days to 50% flowering, plant height, and panicle length and a significant positive correlation with panicle circumference. It is observed that both these micronutrients are not correlated with grain yield (Table 6).

DISCUSSION

Micronutrient malnutrition is a major concern in West and Central Africa. Breeding OPVs with grain Fe and Zn micronutrients will be helpful in combating the dietary-based malnutrition among the poor households in this region. Significant differences for grain Fe and Zn densities and grain yield indicate the presence of variability for these traits in the material tested. Environment is one of the major factors that influences the performance of the genotypes. As $G \times E$ interaction is important for any trait breeding, ANOVA revealed a higher magnitude of mean squares for genotypes than the mean squares of $G \times E$ interaction, marking the lower impact of the environment on the inheritance of these agronomic and micronutrient traits studied across locations.

In the Sahelian zone of West Africa, the farmers prefer to have the early maturing genotypes with long panicles and dual-purpose pearl millet where the rainfall is 300–400 mm (Blümmel et al., 2003; Omany et al., 2007; Anna Ida Pucher, 2018). In this study, most of the genotypes exhibited early flowering, which is very important for drought-prone areas of Sub-Saharan Africa,

especially in the Sahelian zone of West Africa where millet is grown in around 8 million ha out of 16 million ha of total millet cultivation in Africa (Zakari et al., 2019).

CHAKTI and ICMP 177002 have performed well with high grain yield and grain Fe and Zn contents. CHAKTI has iniadi background, which was developed using the intrapopulation method after four cycles of recurrent selections for high Fe. CHAKTI is proposed for the large-scale farmer trials and release in WCA. Previous studies reported that iniadi germplasm has been found to be the best source for grain Fe and Zn contents (Velu et al., 2011; Rai et al., 2014; Govindaraj et al., 2020). The identification of new sources of grain Fe and Zn is equally important as depending on one type of germplasm may be counterproductive in achieving micronutrient genetic gains in pearl millet.

The AMMI ANOVA also indicated the broad range of diversity that existed among genotypes, and the significance of environments and their interactions indicated that there was a differential response of genotypes to environments and will help breeder in choosing genotypes to a specific environment. The significance of IPCA1 and IPCA2 indicated the presence of a complex, multidimensional variation in the $G \times E$ data. The AMMI biplot analysis is a predominant method to find the $G \times E$ interaction for different studied traits. Low $G \times E$ interaction indicates the stability of the genotypes over the range of environments. A genotype showing high positive interaction in an environment can exploit the agroecological or agromanagement conditions of the specific environment and is, therefore, best suited to that environment. Heritability across the locations is very significant indicating a relatively lesser $G \times E$ influence; thus, simple selection for micronutrients is highly feasible. These magnitudes of heritability are affected by the type of genetic material, traits variability, and environmental conditions to which the trials are conducted (Dabholkar, 1992; Falconer and Mackay, 1996). Therefore, the heritability estimate is relatively high for grain yield due to conducive rainfall and other environmental conditions such as the adaptability of the new variety.

The most powerful interpretive tool in the analysis of $G \times E$ interaction in the AMMI model is the biplot analysis. The biplots permit easy visualization of differences in interaction effects. In the AMMI biplot, the genotypes with high mean performance and large value of IPCA scores are conceived as specific adaptability to the environment. Purchase et al. (2000) proposed an ASV measure to quantify and classify genotypes according to their *per se* potential. The ASV has identified the genotype CHAKTI as the most prominent genotype that outperformed across the majority of the locations, indicating this as a stable performing genotype across environments. The ASV parameter has been successfully used in several studies to find stable performers (Mallikarjuna et al., 2015; Govindaraj et al., 2018). CHAKTI is the top-ranking OPV that had a lower ASV of 0.5–2.0 for grain yield and micronutrients, and still, it ranked first in 6 out of 10 locations and second in only 1 location. Based on the multilocal testing and superior performance, recently in 2018, the first biofortified pearl millet variety “CHAKTI” in Africa was released in Niger (Jayashree and Meyer, 2018) and

TABLE 4 | Interaction principal component axis scores of agronomic and morphological traits of pearl millet genotypes across locations in West Africa.

Genotype	Grain yield								Grain Fe content				Grain Zn content			
	Mean	Rank (X)	IPCA 1	IPCA 2	ASV	ASV rank (Y)	YSI (X + Y)	YSI rank	IPCA 1	IPCA 2	ASV	ASV rank	IPCA 1	IPCA 2	ASV	ASV rank
AFRIBEH-NAARA	0.79	19	0.05	−0.11	0.19	3	22	6	1.68	0.88	2.14	28	1.10	−0.01	1.66	18
AKAD-KOM	0.83	17	0.09	0.05	0.26	7	24	8	−1.35	0.27	1.59	26	−0.59	0.00	0.89	7
CHAKTI	1.42	1	0.18	0.16	0.54	15	16	4	0.82	0.21	0.98	16	1.39	−0.36	2.14	25
Faringuero	0.73	21	−0.26	0.08	0.76	18	39	18	−0.15	0.49	0.52	6	−1.43	0.20	2.18	26
Gamoji	1.19	7	−0.05	−0.20	0.25	5	12	2	−0.29	−0.04	0.34	4	−1.02	−0.67	1.68	19
GB 8735	0.83	17	0.05	−0.20	0.25	6	23	7	−0.46	0.48	0.72	10	0.32	0.16	0.50	2
IBMV 8402	0.99	10	−0.59	−0.02	1.71	29	39	18	0.10	0.58	0.59	8	−0.47	−0.92	1.16	14
ICMP 177001	0.85	15	0.28	−0.30	0.86	22	37	16	0.22	−1.38	1.41	22	2.06	0.12	3.11	29
ICMP 177002	1.19	7	0.58	0.25	1.72	30	37	16	−0.20	−0.20	0.31	3	0.69	−1.69	1.99	23
ICMV 167005	1.31	5	−0.19	0.24	0.61	17	22	6	0.32	−0.01	0.37	5	−0.88	0.79	1.55	17
ICMV 167006	1.13	8	−0.29	0.34	0.92	23	31	13	0.89	0.37	1.10	18	0.60	1.86	2.06	24
ICMV 177001	0.61	23	0.14	0.12	0.44	12	35	15	−3.24	1.46	4.04	30	−2.08	1.23	3.38	30
ICMV 177002	0.83	17	0.09	−0.10	0.28	8	25	9	−1.56	−2.19	2.85	29	−0.48	−0.99	1.23	15
ICMV 177003	0.83	17	0.17	−0.21	0.54	16	33	14	0.61	−0.29	0.77	12	0.47	1.58	1.73	20
ICMV 177004	0.80	18	−0.10	0.11	0.32	10	28	11	1.19	−0.26	1.41	23	0.37	0.46	0.73	4
ICMV 221	0.79	19	0.31	−0.14	0.92	24	43	20	−0.01	−0.05	0.05	1	0.33	0.78	0.93	9
ICMV 167001	1.33	3	0.27	0.25	0.83	20	23	7	0.16	0.70	0.72	11	−1.92	−0.74	2.99	28
ICMV IS 85327	0.97	12	−0.29	−0.41	0.94	26	38	17	1.01	−0.10	1.17	19	0.41	−0.08	0.63	3
ICMV IS 94222	1.28	6	−0.17	0.17	0.52	14	20	5	0.41	0.21	0.52	7	−0.24	1.10	1.16	13
ICMV IS 99001	1.32	4	−0.10	−0.01	0.30	9	13	3	0.15	0.15	0.23	2	−0.59	−0.37	0.97	11
ICTP 8203	0.75	20	0.11	0.02	0.33	11	31	13	0.40	−0.69	0.83	13	1.49	0.68	2.35	27
Jirani	0.84	16	0.08	−0.09	0.25	4	20	5	−0.25	−1.20	1.24	21	0.58	0.33	0.93	10
KAANATI	0.69	22	0.37	−0.19	1.08	27	49	22	−0.50	1.35	1.47	24	0.13	−0.79	0.82	5
LCIC 9702	0.87	14	−0.03	−0.10	0.13	2	16	4	0.20	−0.86	0.89	14	1.19	−0.48	1.86	22
Moro	1.06	9	−0.25	−0.40	0.84	21	30	12	0.19	−0.97	1.00	17	0.38	−1.65	1.75	21
Sadore Local	0.73	21	−0.32	0.15	0.93	25	46	21	−0.47	0.46	0.71	9	−0.79	−0.39	1.25	16
SARIA-1	0.61	23	0.27	0.16	0.79	19	42	19	−0.98	−1.79	2.13	27	0.22	−0.86	0.92	8
SOSAT-C88	1.38	2	0.01	0.01	0.04	1	3	1	0.46	0.75	0.92	15	−0.76	0.10	1.15	12
Souna 3	0.98	11	−0.55	0.13	1.62	28	39	18	1.00	0.19	1.19	20	−0.31	0.76	0.89	6
WAAPP-NAARA	0.88	13	0.15	0.26	0.50	13	26	10	−0.36	1.49	1.55	25	−0.15	−0.16	0.28	1

IPCA 1, interaction principal component axis one score; IPCA 2, interaction principal component axis two score; ASV, AMMI Stability Value; YSI, yield stability index.

TABLE 5 | First four AMMI selections based on best grain yielding genotypes in each environment.

Locations	Grain yield (t/ha)	Effect	Rank			
			1	2	3	4
Maradi (Niger)	0.94	−0.6154	ICMV IS 99001	SOSAT-C88	Moro	IBMV 8402
Magaria (Niger)	0.98	−0.5197	ICMV 167005	ICMV IS 94222	SOSAT-C88	ICMV IS 99001
Bambey (Senegal)	0.96	0.0596	CHAKTI	SOSAT-C88	ICMV IS 99001	ICMV 167001
Nioro (Senegal)	1.07	0.367	CHAKTI	ICMV 167001	ICMP 177002	SOSAT-C88
Sadore (Niger)	0.90	−0.2665	ICMV 167005	CHAKTI	ICMV IS 94222	ICMV 167001
Kano (Nigeria)	0.70	0.1276	CHAKTI	ICMV 167001	ICMV 167005	SOSAT-C88
Konni (Niger)	0.85	0.4307	CHAKTI	ICMV 167001	ICMP 177002	SOSAT-C88
Bawku (Ghana)	0.94	0.5633	CHAKTI	SOSAT-C88	ICMP 177002	ICMV 167001
Gampela (Burkina Faso)	1.06	0.5125	CHAKTI	ICMP 177002	ICMV 167001	SOSAT-C88
Cinzana (Mali)	1.20	−0.6591	ICMV 167005	ICMV IS 99001	ICMV 167001	SOSAT-C88
Best candidate variety			CHAKTI	SOSAT-C88 & ICMV167001	ICMV 167001 & ICMP177002	SOSAT-C88

TABLE 6 | Association of agronomic and morphological traits of pearl millet (*Pennisetum glaucum* L.) across locations in West Africa.

Traits	Downy mildew damage%	Days to 50% flowering	Plant height	Panicle length	Panicle circumference	Grain yield	Grain Fe content	Grain Zn content
Downy mildew damage%	1.00							
Days to 50% flowering	−0.38*	1.00						
Plant height	−0.44*	0.93**	1.00					
Panicle length	−0.20	0.68**	0.68**	1.00				
Panicle circumference	0.31	−0.62**	−0.68**	−0.58**	1.00			
Grain yield	−0.13	0.02	0.08	0.33	0.10	1.00		
Grain Fe content	0.45*	−0.77**	−0.74**	−0.70**	0.55**	0.01	1.00	
Grain Zn content	0.39*	−0.77**	−0.75**	−0.70**	0.57**	0.01	0.98**	1.00

*, **Significant at $P = 0.05$ and 0.01 , respectively.

the ECOWAS regional catalog in 2019¹. This variety is gaining importance in the farming community in Niger, Mali, Burkina Faso, and Senegal due to its extra early maturity, high yield, and high grain Fe and Zn contents (Mavindidze, 2020).

Association studies among traits revealed the presence of a significant positive correlation between the grain Fe and Zn contents, indicating that they are inherited together and both traits can be selected in every generation. Similar results were earlier reported in pearl millet (Govindaraj et al., 2013; Anand et al., 2014; Govindaraj et al., 2020). This high correlation may be due to the overlapping and co-segregating quantitative trait loci reported in pearl millet (Kumar, 2011), implying that selecting high-Fe lines would likely increase Zn as an associated trait. Grain Fe and Zn contents have exhibited no significant correlation with grain yield, which implies that there is no yield limitation in increasing micronutrients in pearl millet. Earlier studies in pearl millet reported a significant negative correlation with grain yield (Rai et al., 2012; Anand et al., 2014) and no correlations (Govindaraj et al., 2019; Pujar et al., 2020), indicating that conscious selection for both traits is required in segregating population to avoid any negative linkages. Also, it is important to develop the experimental population involving diverse high-yield and high-Fe/Zn parents to study the linkage between

yield and nutrient traits in the segregating population or by using genomic tools.

CONCLUSION

The biofortification in pearl millet intends to contribute towards the prevention of micronutrient deficiencies by increasing the daily adequacy of micronutrient intakes among the poor farming population in the semi-arid regions. A stable performing high-Fe variety “CHAKTI” was identified across 10 locations, with high grain yield (71% higher) and grain Fe (16% higher) and Zn (7.4% higher) contents than the currently growing OPVs (35 ppm mean Fe and 28 ppm mean Zn). The correlation between Fe and Zn and no correlation with yield indicate that both micronutrients can be improved simultaneously in high-yielding backgrounds. Further study is needed to explore the progeny test-based selections from the identified high-Fe/Zn OPVs for breeding hybrid parents prospecting biofortification in WCA with competitive yield in near future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

¹<http://www.coraf.org/paired/wp-content/uploads/2019/11/Catalogue-Re%CC%81gional-des-Espe%CC%80ces.pdf>

AUTHOR CONTRIBUTIONS

PG and MG conceived the concept and provided the genetic materials. MR and PG helped in designing the trial, data curation, conducted the research, statistical data analyses, and drafting the manuscript. DI conducted trials in Burkina Faso and contributed to the data analysis and data curation. OS conducted trials in Senegal and contributed to the data curation and data analysis from Senegal. KAI helped in conducting multisite data from Niger. PAA contributed to conducting the trial and data curation from the Ghana trial. MDS helped in conducting the trial and data curation from Mali locations. AII conducted trials in Nigeria and contributed to the data curation from Nigeria. All authors read and agreed to the content of the manuscript.

REFERENCES

- Anand, K., Rai, K. N., Radhika, K., Govindaraj, M., Sahrawat, K. L., and Rao, A. S. (2014). Grain iron and zinc density in pearl millet: combining ability, heterosis and association with grain yield and grain size. *Springerplus* 3:763. doi: 10.1186/2193-1801-3-763
- Anna Ida Pucher (2018). *Pearl Millet Breeding in West Africa – Steps Towards Higher Productivity and Nutritional Value. Dissertation Submitted in Fulfilment of the Requirements for the Degree.* Germany: Universität Hohenheim.
- Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., et al. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 382, 427–451. doi: 10.1016/S0140-6736(13)60937-X
- Blümmel, M., Zerbini, E., Reddy, B. V. S., Charles Thomas, H. Jr., Bidinger, F., Khan, A. A., et al. (2003). Improving the production and utilization of sorghum and pearl millet as livestock feed: progress towards dual-purpose genotypes. *F. Crop Res.* 84, 143–158. doi: 10.1016/S0378-4290(03)00146-1
- Bouis, H. E., Hotz, C., McClafferty, B., Meenakshi, J. V., and Pfeiffer, W. H. (2011). Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 32, S31–S40.
- Dabholkar, A. R. (1992). *Elements of Biometrical Genetics.* New Delhi: Concept Publishing Company, 431.
- Falconer, D. S., and Mackay, T. F. C. (1996). *Quantitative Genetics*, 4th Edn. London: Longman Group Limited, 464.
- Finkelstein, J. L., Mehta, S., Udipi, S. A., Ghugre, P. S., Luna, S. V., Wenger, M. J., et al. (2015). A randomized trial of iron-biofortified pearl millet in school children in India. *J. Nutr.* 145, 1576–1581. doi: 10.3945/jn.114.208009
- GenStat (2015). *Genstat for Windows*, 18th Edn. Hemel Hempstead: VSN International.
- Govindaraj, M., Rai, K. N., Kanatti, A., Rao, A. S., and Shivade, H. (2019). Nutritional security in drylands: fast-track intra-population genetic improvement for grain iron and zinc densities in pearl millet. *Front. Nutr.* 6:74. doi: 10.3389/fnut.2019.00074
- Govindaraj, M., Rai, K. N., Kanatti, A., and Shivade, H. (2018). Terminal drought and a d 2 dwarfing gene affecting grain iron and zinc concentration in pearl millet. *J. Cereal Sci. (TSI)* 79, 247–252. doi: 10.1016/j.jcs.2017.11.005
- Govindaraj, M., Rai, K. N., Kanatti, A., Upadhyaya, H. D., Shivade, H., and Rao, A. S. (2020). Exploring the genetic variability and diversity of pearl millet core collection germplasm for grain nutritional traits improvement. *Sci. Rep.* 10:21177.
- Govindaraj, M., Rai, K. N., Shanmugasundaram, P., Dwivedi, S. L., Sahrawat, K. L., Muthaiah, A. R., et al. (2013). Combining ability and heterosis for grain iron and zinc densities in pearl millet. *Crop Sci.* 53, 507–517. doi: 10.2135/cropsci2012.08.0477
- Guild, G. E., and Stangoulis, J. C. R. (2016). Non-matrix matched glass disk calibration standards improve XRF micronutrient analysis of wheat grain

ACKNOWLEDGMENTS

We are thankful to the ICRISAT pearl millet breeding team Hamadou Adamou, Issa Karimou, Tahirou Boye, and Ada Abarachi based in Niamey, Niger for their help in setting the trial, and carrying out the field experiments at Sadore, Konni, and Magaria stations. We are also thankful to ICRISAT, Kano, Nigeria for conducting trials in Nigeria. NARS partners from INRAN, Niger; INERA, Burkina Faso; ISRA, Senegal; CSIR-SARI, Ghana; and IER, Mali who have helped us in multilocation trials and the release committee, Government of Niger for their efforts in the release of “CHAKTI.” Special thanks to millet team members. We also thank the CGIAR HarvestPlus program, Bill and Melinda Gates Foundation, and CRP Grain Legumes and Dryland Cereals for providing the financial support.

- across five Laboratories in India. *Front. Plant Sci.* 7:784. doi: 10.3389/fpls.2016.00784
- Haas, J. E. (2014). Efficacy and other nutrition evidence for iron crops. *Biofortification Prog. Briefs*, 20, Washington, D.C.: International Food Policy Research Institute (IFPRI). 39–40.
- ICRISAT archives Available online at: <http://www.icrisat.org/what-we-do/crops/PigeonPea/Archives/Earlymillet.htm>
- Jayashree, B., and Meyer, C. (2018). *Africa's First Biofortified Pearl Millet Variety Aims to Combat Anaemia.* Available online at: <https://www.icrisat.org/africas-first-biofortified-pearl-millet-variety-aims-to-combat-anaemia/> (accessed July 19, 2020).
- Kodkany, B. S., Bellad, R. M., Mahantshetti, N. S., Westcott, J. E., Krebs, N. F., Kemp, J. F., et al. (2013). Biofortification of pearl millet with iron and zinc in a randomized controlled trial increases absorption of these minerals above physiologic requirements in young children. *J. Nutr.* 143, 1489–1493. doi: 10.3945/jn.113.176677
- Kramer, C. V., and Allen, S. (2015). Malnutrition in developing countries. *Paediatr. Child Health* 25, 422–427.
- Kumar, S. (2011). *Development of New Mapping Population and Marker-assisted Improvement of Iron and Zinc Grain Density in Pearl Millet [Pennisetum glaucum (L.) R. Br.]*. Dissertation, Bikaner: Swami Keshwanand Rajasthan Agricultural University.
- Mallickarjuna, M. G., Thirunavukkarasu, N., Hossain, F., Bhat, J. S., Jha, S. K., Rathore, A., et al. (2015). Stability performance of inductively coupled plasma mass spectrometry-phenotyped kernel minerals concentration and grain yield in maize in different agro-climatic zones. *PLoS One* 10:e0139067. doi: 10.1371/journal.pone.0139067
- Mavindidze, D. (2020). *Iron Pearl Millet: Crop of the Future for Improving Nutrition in Sahel.* Available online at: <https://www.harvestplus.org/knowledge-market/in-the-news/iron-pearl-millet-crop-future-improving-nutrition-sahel> (accessed July 27, 2020).
- Omanya, G. O., Weltzien-Rattunde, E., Sogodogo, D., Hanssens, N., Guero, Y., Zangre, R. G., et al. (2007). Participatory varietal selection with improved pearl millet in West Africa. *Exp. Agric.* 43, 5–19. doi: 10.1017/S0014479706004248
- Paltridge, N. G., Milham, P. J., Ortiz-Monasterio, J. I., Velu, G., Yasmin, Z., et al. (2012a). Energy-dispersive X-ray fluorescence spectrometry as a tool for zinc, iron and selenium analysis in whole grain wheat. *Plant Soil* 361, 261–269. doi: 10.1007/s11104-012-1423-0
- Paltridge, N. G., Palmer, L. J., Milham, P. J., Guild, G. E., and Stangoulis, J. C. R. (2012b). Energy-dispersive X-ray fluorescence analysis of zinc and iron concentration in rice and pearl millet grain. *Plant Soil* 361, 251–260. doi: 10.1007/s11104-011-1104-4
- Pujar, M., Govindaraj, M., Gangaprasad, S., Kanatti, A., and Shivade, H. (2020). Genetic variation and diversity for grain iron, zinc, protein and agronomic traits in advanced breeding lines of pearl millet [*Pennisetum glaucum* (L.) R. Br.] for biofortification breeding. *Genet. Resour. Crop. Evol.* 67, 2009–2022. doi: 10.1007/s10722-020-00956-x

- Purchase, J. L., Hatting, H., and van Deventer, C. S. (2000). Genotype x environment interaction of winter wheat (*Triticum aestivum* L.) in South Africa: II. Stability analysis of yield performance. *South Afr. J. Plant Soil* 17, 101–107.
- Rai, K. N., Govindaraj, M., and Rao, A. S. (2012). Genetic enhancement of grain iron and zinc content in pearl millet. *Qual. Assur. Saf. Crop.* 4, 119–125. doi: 10.1111/j.1757-837X.2012.00135.x
- Rai, K. N., Gupta, S. K., Sharma, R., Govindaraj, M., Rao, A. S., Shivade, H., et al. (2014). Pearl millet breeding lines developed at ICRISAT: a reservoir of variability and useful source of non-target traits. *SAT e-J.* 1, 1–13.
- Rai, K. N., Kumar, K. A., Andrews, D. J., Rao, A. S., Raj, A. G. B., and Witcombe, J. R. (1990). Registration of 'ICTP 8203' pearl millet. *Crop Sci.* 30:959. doi: 10.2135/cropsci1990.0011183x003000040048x
- Saltzman, A., Birol, B., Bouis, H. E., Boy, E., De Moura, F. F., Islam, Y., et al. (2013). Biofortification: progress toward a more nourishing future. *Glob Food Secur.* 2, 9–17. doi: 10.1016/j.gfs.2012.12.003
- Saltzman, A., Birol, E., Oparinde, A., Andersson, M. S., Asare-Marfo, D., Diressie, M. T., et al. (2017). Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. *Ann. N. Y. Acad. Sci.* 1390, 104–114. doi: 10.1111/nyas.13314
- Tulchinsky, T. H. (2010). Micronutrient deficiency conditions: global health issues. *Public Health Rev.* 32, 243–255. doi: 10.1007/BF03391600
- Underwood, B. A. (2000). Overcoming micronutrient deficiencies in developing countries: is there a role for agriculture? *Food Nutr. Bull.* 21, 356–360.
- UNICEF (2019). *The State of the World's Children 2019. Children, Food and Nutrition: Growing Well in a Changing World*. New York, NY: UNICEF.
- Velu, G., Ortiz-Monasterio, I., Singh, R. P., and Payne, T. (2011). Variation for grain micronutrients in wheat core collections accession of diverse origin. *Asian J. Crop Sci.* 3, 43–48. doi: 10.3923/ajcs.2011.43.48
- Webb, P., Stordalen, G. A., Singh, S., Wijesinha-Bettoni, R., Shetty, P., and Lartey, A. (2018). Hunger and malnutrition in the 21st century. *BMJ* 361:k2238. doi: 10.1136/bmj.k2238
- World Health Organization (2006). *Workshop to Review the Results of Studies Evaluating the Impact of Zinc Supplementation on Childhood Mortality and Severe Morbidity: Conclusions and Next Steps*. Geneva: World Health Organization.
- Zakari, H., Mohammed, R., Gangashetty, P. I., Govindaraj, M., and Tangara, M. (2019). Improvement of restorer lines for strengthening pearl millet (*Pennisetum glaucum* L.) hybrid breeding in West and Central Africa. *J. Agric. Crop Res.* 7, 204–214.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Gangashetty, Riyazaddin, Sanogo, Inousa, Issoufou, Asungre, Sy, Govindaraj and Ignatius. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Understanding Heterosis, Genetic Effects, and Genome Wide Associations for Forage Quantity and Quality Traits in Multi-Cut Pearl Millet

Ponnaiah Govintharaj^{1,2*}, Marappa Maheswaran¹, Michael Blümmel³, Pichaikannu Sumathi¹, Anil Kumar Vemula², Abhishek Rathore², Selvanayagam Sivasubramani², Sandip Mallikarjun Kale², Rajeev Kumar Varshney² and Shashi Kumar Gupta²

OPEN ACCESS

Edited by:

Palak Chaturvedi,
University of Vienna, Austria

Reviewed by:

Yong Suk Chung,
Jeju National University, South Korea
Mehathan Muthamilarasan,
University of Hyderabad, India
Arindam Ghatak,
University of Vienna, Austria

*Correspondence:

Ponnaiah Govintharaj
p.govintharaj@cgia.org;
govintha.tnau@gmail.com

Specialty section:

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

Received: 30 March 2021

Accepted: 14 October 2021

Published: 18 November 2021

Citation:

Govintharaj P, Maheswaran M, Blümmel M, Sumathi P, Vemula AK, Rathore A, Sivasubramani S, Kale SM, Varshney RK and Gupta SK (2021) Understanding Heterosis, Genetic Effects, and Genome Wide Associations for Forage Quantity and Quality Traits in Multi-Cut Pearl Millet. *Front. Plant Sci.* 12:687859. doi: 10.3389/fpls.2021.687859

¹ Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, India, ² International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India, ³ International Livestock Research Institute (ILRI), Patancheru, India

Pearl millet is an important food and fodder crop cultivated in the arid and semi-arid regions of Africa and Asia, and is now expanding to other regions for forage purpose. This study was conducted to better understand the forage quantity and quality traits to enhance the feed value of this crop. Two sets of pearl millet hybrids (80 single cross hybrids in Set-I and 50 top cross hybrids in Set-II) along with their parents evaluated multi-locationally for the forage-linked traits under multi-cut (two cuts) system revealed significant variability for the forage traits in the hybrids and parents. The mean better parent heterosis (BPH) for total dry forage yield (TDFY) was 136% across all the single cross hybrids and 57% across all the top cross hybrids. The mean BPH for *in vitro* organic matter digestibility (IVOMD) varied from –11 to 7% in the single cross hybrids and –13 to 11% in the top cross hybrids across cuts. The findings of TDFY and IVOMD heterosis in these sets indicated the potential of improvement of the hybrid cultivars for forage quantity and quality in forage pearl millet. The parental lines single cross parent (SCP)-L02, SCP-L06, and top cross parent (TCP)-T08 found superior in the forage quantity and quality traits can be utilized in the future breeding programs. Most of the forage traits were found to be controlled by using the non-additive gene action. A diverse panel of 105 forage-type hybrid parents (Set-III) genotyped following genotyping by sequencing (GBS) and phenotyped for crude protein (CP) and IVOMD under multi-cuts for 2 years identified one stable significant single nucleotide polymorphism (SNP) on LG4 for CP, and nine SNPs for IVOMD distributed across all the linkage groups except on LG2. The identified loci, once validated, then could be used for the forage quality traits improvement in pearl millet through marker-assisted selection.

Keywords: single cross hybrids, top cross hybrids, line × tester, general and specific combining ability, non-additive gene action, association mapping, gene annotation, biomass

INTRODUCTION

Pearl millet [*Pennisetum glaucum* (L) R.Br.] is cultivated mainly in the hot and dry agro-ecologies of Africa and Asia. This crop has a high potential for biomass production due to its C_4 photosynthetic pathway, and it additionally possesses tolerance to various climatic stresses. Currently, this crop is gaining popularity among the small-holder farmers because of its potential to adapt to the diverse agro-climatic conditions where most of the other crops, such as wheat, rice, sorghum, maize, and barley fail to produce economic yields (Baltensperger, 2002; Vadez et al., 2012).

Pearl millet is primarily cultivated for human consumption in the developing countries. Besides grain production, it is grown as feed and forage crop for livestock grazing, silage, hay, and green fodder chopping (Newman et al., 2010) and is fed to the animals at any crop growth stage without any adverse effect (Arya et al., 2013) in a range of countries, such as the United States (Sheahan, 2014), in the summer season in Australia (Hanna, 1996), Canada (Brunette et al., 2014), Mexico (Urrutia et al., 2015), in triple cropping system in southern Kyushu, Japan (Li et al., 2019), Iran (Aghaalkhani et al., 2008), Central Asia (Nurbekov et al., 2013), and Brazil (Dias-Martins et al., 2018). Recently, it has emerged as an important fodder crop during the summer months of north-western India (Amarendra Reddy et al., 2013).

Lack of sufficient quantity of fodder is the major constrain to livestock production in the smaller farming communities in the arid and semi-arid regions. For instance, at present the world has feed shortage of about 911 million tons and would require 1,148 million tons feed by 2030 (Food and Agriculture Organization [FAO], 2002). To alleviate the feed shortage in these regions, further exploitation of pearl millet could be one of the promising solutions as it is well adapted to the arid and semi-arid regions with several other benefits, such as high tillering potential and quick regenerative ability assuring the possibility of multi-cutting, which allows the year-round supply of green/dry forage (Babiker et al., 2014). The results from the multi-location trials conducted during the summer season under the All India Coordinated Forage Project showed pearl millet varieties had higher green forage yield (GFY: 38 t ha⁻¹) and crude protein (CP: 9%) in comparison with sorghum (33 t ha⁻¹

and 6%) and maize (31 t ha⁻¹ and 5.5%) varieties in single cut (Rai et al., 2004). Almost all the released forage cultivars of pearl millet available in the market till date are based on single cut, but milk producing farmers in the semi-arid regions are now demanding multi-cut (2–3 cuts/season) forage with increased digestibility to meet year-round supply of forage/feed. In addition, the pearl millet cultivars bred in the past were dual purpose types with no focus on the improved forage quality, hence were rejected by the farmers because of poor fodder quality of the stover (Kelley et al., 1996). The need has been felt to breed exclusive forage multi-cut type cultivars with better quality to feed the livestock.

The studies based on the animal experiments concluded that the cows fed with pearl millet silage produced milk with increased milk fat concentration than those fed with corn silage (Amer and Mustafa, 2010). Similarly, the cows fed with pearl millet silage consumed more dry matter than those fed with sorghum silage or tropical corn silage (Amer and Mustafa, 2010). Most recently, Lauriault et al. (2021) observed that both pearl millet and sorghum-sudangrass produced equal dry matter yield at single cut forage, and the pearl millet crop provided greater average daily gains than sorghum-sudangrass in beef when both have the same levels of nutritive value. In another livestock performance study by Vinutha et al. (2021), Nellore ram lambs fed with pearl millet silage showed increased digestibility than those fed with sorghum silage harvested at 76 days after sowing. These studies suggest that pearl millet is an excellent choice for the farmers and milch animals in terms of forage yield and quality as compared with other cereals in the drier regions of the world. Hence, targeted breeding efforts have been initiated in last one decade to improve the forage yield and quality traits in pearl millet (Rai et al., 2012; Gupta et al., 2015; Ponnaiah et al., 2019).

Exploitation of heterosis in pearl millet is considered as an easy tool with its protogynous nature of flowering and availability of stable cytoplasmic male sterility (CMS) system (Burton, 1965). Several studies have been conducted to assess the hybrid performance and heterosis in the grain type parents/populations in pearl millet (Pucher et al., 2016; Gupta et al., 2018, 2020; Sattler et al., 2019; Singh and Gupta, 2019; Patil et al., 2020; Sattler and Haussmann, 2020; Dutta et al., 2021). No research has been undertaken to assess the magnitude of heterosis exclusively for the forage yield and quality traits in multi-cut pearl millet.

Limited number of quantitative trait loci (QTL) mapping studies have been conducted in pearl millet for the fodder linked traits, such as for tiller numbers (Varshney et al., 2017), stover yield (Yadav et al., 2002; Nepolean et al., 2006), *in vitro* organic matter digestibility (IVOMD), and nitrogen content (Nepolean et al., 2006). Hash et al. (2003) reported that the QTL mapping and marker-assisted selection (MAS) can be implied for stover yield, forage disease resistance, and for *in vitro* estimates of the nutritive value of several stover fractions of pearl millet for ruminants.

Considering that there is limited information on the gene action, heterosis, and mapping of forage traits under multi-cut system in forage type pearl millet, the present study aimed to estimate heterosis, general combining ability (GCA), and

Abbreviations: SCH, single cross hybrid; TCH, top cross hybrid; SCP, single cross parent; TCP, top cross parent; PH, plant height; GFY, green forage yield; TGFY, total green forage yield; DFY, dry forage yield; TDFY, total dry forage yield; NIRS, near-infrared reflectance spectroscopy; SCP, crude protein; NDF, neutral detergent fiber; ADF, acid detergent fiber; ADL, acid detergent lignin; ME, metabolizable energy; IVOMD, *in vitro* organic matter digestibility; FC, first cut; SC, second cut; BP, better parent; BPH, better parent heterosis; CH, check hybrid; SH, standard heterosis; OPV, open pollinated variety; SCA, specific combining ability; GCA, general combining ability; GBS, genotyping by sequencing; PCR, polymerase chain reaction; MAF, minor allele frequency; SNP, single nucleotide polymorphism; PCA, principal component analysis; BLUP, best linear unbiased predictors; TASSEL, Trait Analysis by aSSociation, Evolution and Linkage; GAPIT, Genome Association and Prediction Integrated Tool; MLM, multi-locus mixed model; QQ, quantile-quantile; MAS, marker assisted selection; QTL, quantitative trait loci; CMS, cytoplasmic male sterility; IVDM, *in vitro* dry matter digestibility; ICRIAT, International Crops Research Institute for the Semi-Arid Tropics; TNAU, Tamil Nadu Agricultural University; PR, predictability ratio; ANOVA, analysis of variance.

specific combining ability (SCA) for the forage traits of single cross and top cross hybrids at two cutting intervals, and to identify the QTLs for the important forage quality traits in pearl millet.

MATERIALS AND METHODS

Plant Materials

In the present study, two sets of test crosses were developed to investigate heterosis, combining ability, and gene effects for the forage traits. The first set consisted of 80 single cross hybrids (hereafter, referred as Set-I) generated by crossing 10 seed (A-lines) parents and eight pollinator (R-lines) parents having forage type traits (**Supplementary Table 1**) in a line \times tester mating design. The second set comprised of 50 top cross hybrids (hereafter, referred as Set-II) developed by crossing five seed parents (A-lines) and 10 germplasm accessions/ open pollinated varieties (OPVs) as pollinators (**Supplementary Table 2**) in line \times tester mating design. These pollinators were already identified as the promising breeding lines for having higher biomass (Gupta et al., 2015).

A third set (Set-III) of 105 diverse hybrid parents comprised of 17 seed and 88 pollinator parents [such as parental lines from Set-I (one pollinator parent not included due to poor quality of DNA from Set-I) and Set-II (excluding pollinators of set-II which were populations)] (**Supplementary Table 3**) derived from an advanced high biomass nursery (F_6 and above) of the pearl millet breeding program of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana, India, was used to examine for the genome wide association study (GWAS) for the two forage quality traits CP and IVOMD.

Evaluation of Parental and Hybrid Trials

Both the single cross (Set-I) and top cross (Set-II) hybrids were evaluated along with the parents and commercial checks at two locations, namely, ICRISAT, Patancheru (18°N, 78°E, and 545 m above sea level) and Tamil Nadu Agricultural University (TNAU), Coimbatore, Tamil Nadu, India (11°N, 77°E, and 411.98 m above sea level) during the summer season of 2015. The hybrids and parents were planted side by side in each trial as separate blocks in each replication to avoid any suppressive effect of the hybrids over parents. The parental block included the fertile (B-lines) lines corresponding to their male sterile (A-lines) lines used in the development of hybrids. Four forage pearl millet commercial hybrids popular in India, namely, PAC 981 (Nutrifeed), DFMH 31, Milkon, and Poshan were used as checks in the hybrids block.

Both the hybrids and parents were planted in alfisol soils in alpha lattice design with three replications at ICRISAT, Patancheru. Each entry was planted in the four rows of 4 m length with rows spaced 60 cm apart and the plants spaced at 10–12 cm from each other. All the hybrids and parents were planted in black soils in alpha lattice design with the two replications at TNAU, Coimbatore. Each entry was planted in the four rows, each of 4 m length and the rows spaced 45 cm apart. The experiment was conducted during first week of March to the third week of June

2015. The trial was irrigated at 12–15 days interval to avoid any moisture stress, and the crop was protected from the weeds, pests, diseases, and animals at both the locations.

The third set (Set-III) of 105 hybrid parents was evaluated at ICRISAT, Patancheru, in alfisol soils during the summer seasons of 2015 and 2016 in partially balanced alpha lattice design with two replications. The plot size, spacing, and field management were same as specified above in the evaluation of first and second sets of trials (of hybrids and parents) at ICRISAT. The experiment was conducted during the first week of March to the third week of June 2015 and 2016.

Recording of Morphological and Biochemical Traits

In both Set-I and Set-II trials, the four rows of each entry (hybrids and parents) were harvested manually by cutting at the second node from the bottom of the plant at 50–55 days after planting (at boot stage of plants) as the first cut. At the time of harvest, the plant height (PH, centimeters) was measured on five random plants from the base of the stem to the tip of the panicle of the main tiller. Fresh weight of the green forage was recorded (kilogram) for each plot. A sub-sample (10–15 plants) of about 1 kg was collected per entry at the time of harvest and recorded for green forage weight, oven dried for 8 h daily for 3–4 days at 60°C in Campbell dryer (Campbell Industries, Inc., 3201 Dean Avenue, Des Moines, IA, United States), and weighed again (dry forage weight in kilogram). The dry matter concentration and dry forage yield (DFY) of each entry was calculated using the formula:

Dry matter concentration (DMC)

$$= \text{Dry forage weight} / \text{Green forage weight} \times 100$$

$$\text{DFY} = \text{Green forage weight} \times \text{dry matter concentration}$$

The GFY and DFY were expressed in units of tons per hectare (t ha^{-1}). The dried sub-samples of the whole plants (10–15 plants) of each entry were chopped into 10–15 mm pieces using a chaff cutter (Model # 230, Jyoti Ltd. Vadodara-India) and ground using Thomas Wiley mill (Model # 4, Philadelphia, PA, United States) to pass through 1-mm screen for the chemical analysis. The ground stover samples (approximately, 40 g of sample/entry) were analyzed by near-infrared reflectance spectroscopy (NIRS) for stover nitrogen concentration ($\text{N} \times 6.25$ equals to CP content), neutral detergent fiber (NDF), acid detergent fiber (ADF), acid detergent lignin (ADL), metabolizable energy (ME), and IVOMD as described by Bidinger and Blümmel (2007) and Blümmel et al. (2007). The second cut of forage was harvested from the same four rows (which were subjected to first cut) after 30 days. The forage quantity and quality traits were also recorded in the second cut, as similar to how it was done in the first cut. The total green forage yield (TGFY) and total dry forage yield (TDFY) (t ha^{-1}) were estimated as the sum of the two cuts for each entry in this trial.

In case of the Set-III of diverse hybrid parents trial, the same cutting methodology and the measurement of traits (CP and

IVOMD) over 2 years of evaluation were followed as similar to that of first and second sets of evaluation trials at ICRISAT.

DNA Extraction, Genotyping by Sequencing, and Single Nucleotide Polymorphism Calling in the Set-III (Hybrid Parents Panel)

The fresh leaf tissues (100 mg) were collected from 20 to 25 seedlings (per accession) of 8 days old plants (days after sowing) planted in the pots in darkroom at 18°C–25°C, and the DNA was isolated using the NucleoSpin® 96 Plant II kit (Macherey-Nagel, Düren, Germany). Normalization of genomic DNA to 10 ng/μl was done in 0.8% agarose gel using lambda DNA (MBI Fermentas, Hanover, MD, United States). Electrophoresis was performed in Tris acetate-ethylenediaminetetraacetic acid (EDTA) buffer in buffer tank at 90 V for 1 h and the gels were stained with ethidium bromide and visualized in UV light using image analyzer. The DNA of 105 hybrid parents was genotyped using genotyping by sequencing (GBS) (Elshire et al., 2011) by digesting DNA with *ApeKI* endonuclease restriction enzyme. The PCR amplification of pooled amplicons was carried out before sequencing on Illumina HiSeq2500 platform (Illumina Inc, San Diego, CA, United States). The raw sequencing reads and barcode information were processed for single nucleotide polymorphism (SNP) identification from the published pearl millet reference genome (Varshney et al., 2017) using TASSEL v4.0 software (Bradbury et al., 2007). Barcode containing reads were retained and used for SNP calling. These reads were trimmed to 64 bp from barcode side, aligned against each other and used for SNP identification. The identified SNPs were assigned to each hybrid parent based on the information of the barcode sequence. Further, the SNP data were filtered with minor allele frequency (MAF) cut-off of 0.10 (10%) and SNP with $\geq 25\%$ missing data resulted 77,892 SNP markers (Dataset available on ICRISAT Dataverse: <http://dataverse.icrisat.org/dataset.xhtml?persistentId=doi%3A10.21421%2F2FD2%2F1WXCUJ&version=DRAFT>). These were further filtered for site coverage (90%), minor allele frequency (0.05), and maximum heterozygosity (50%), were identified 34,691 SNPs and were used for the GWAS analysis.

Statistical Analysis

The complete parental and hybrids data information were available at first cut of single cross hybrids ($10 \times 8 = 80$ hybrids) and top cross hybrids ($5 \times 10 = 50$ hybrids). In Set-I, the missing data (due to poor regeneration ability of parents and hybrids after first cut) from parents [one line and three testers at one site (TNAU, Coimbatore)] resulted into dropping of 35 single cross hybrids. Hence, $9 \times 5 = 45$ single cross hybrids were analyzed for second cut in Set-I. Whereas, in Set-II, 26 top cross hybrids were removed from further analysis due to the missing data from parents [one line and four testers at one site (TNAU, Coimbatore)] and thus $4 \times 6 = 24$ top cross hybrids were analyzed at second cut.

The genotype means were calculated from the combined ANOVA across the two locations using PROC MIXED

(SAS Institute Inc, 2017). The estimation of better parent heterosis (BPH) and standard heterosis (SH) were calculated using the formula:

$$\text{BPH} = [(F_1 - \text{BP}) / \text{BP}] \times 100, \text{ and,}$$

$$\text{SH} = [(F_1 - \text{CH}) / \text{CH}] \times 100,$$

where, F_1 is the hybrid value, BP (better parent) is either P_1 or P_2 and CH is the value of the check hybrid. BP was considered to have higher values for PH, GFY, TGFY, DFY, TDFY, CP, ME, and IVOMD, and lower values for NDF, ADF, and ADL. Among the four check hybrids, the PAC 981 (Nutrifeed) is a popular high forage yielding hybrid from the Advanta seed company, which occupies large areas of pearl millet for fodder use in India. Therefore, Nutrifeed was considered as a best check hybrid for calculating the SH in the single cross and top cross hybrids.

The line \times tester mating design (Kempthorne, 1957; Arunachalam, 1974) was used to estimate the variance components for Set-I and Set-II hybrids using SAS v9.4 (SAS Institute Inc, 2017). The relative importance of GCA and SCA was also assessed by estimating the components of variance and expressing them in the ratio, $2\sigma^2\text{GCA} / (2\sigma^2\text{GCA} + \sigma^2\text{SCA})$ (Baker, 1978). The Pearson's correlation of hybrid performance with sum of parental GCA effects, SCA effects, and BPH, and also among the BPH of forage traits were determined using GraphPad Prism v5.0 (GraphPad Software, Inc., San Diego, CA, United States).

In GWAS analysis, a set of 105 pearl millet hybrid parents, which were genotyped (34,691 GBS-identified SNPs) and phenotyped (for CP and IVOMD at two cutting intervals) were used for further statistical analysis. The principal component analysis (PCA) was conducted and the resulted eigenvalue (5) was taken into the consideration to correct the population structure along with kinship matrices (K) values implemented in the multi-locus mixed model (MLMM) (Segura et al., 2012) using the R package Genomic Association and Prediction Integrated Tool (GAPIT) (Wang and Zhang, 2020) for GWAS analysis. For population stratification, the Quantile-Quantile (Q-Q) plots were obtained by plots of observed $-\log_{10}$ P -value vs. expected $-\log_{10}$ P -value for all the markers. If there is deviation in the P -values at initial stage, it indicates the existence of population stratification. Manhattan plots were used to visualize the results of GWAS. Negative \log_{10} of the P -values for each SNPs were plotted over the seven chromosomes for the two traits, CP and IVOMD. The best linear unbiased predictors (BLUPs) were estimated for CP and IVOMD for each hybrid parent across years from the combined analysis (SAS Institute Inc, 2017). The Bonferroni correction threshold was used for selecting significant SNP with an alpha value of 0.05. The significant MTA markers were functionally annotated based on the reference pearl millet genome (v1.1) (Varshney et al., 2017) and its features, using SnpEff software version 4.3t (Cingolani et al., 2012). The gene annotations were further mapped to the homologous functions provided at pfam. Broad sense heritability (H^2) of the traits was estimated by

dividing the genotypic variance by the total phenotypic variance (SAS Institute Inc, 2017).

$$H^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2/l + \sigma_e^2/rl),$$

where σ_g^2 is the genotypic variance, σ_{ge}^2 is the genotype \times environment variance, σ_e^2 is the error variance, r represents the replicates, and “ l ” represents the locations.

RESULTS AND DISCUSSION

Variance Analysis for Forage Traits

Combined ANOVA exhibited significant differences for most of the forage traits among the genotypes except TDFY, CP, and IVOMD at second cut in the top cross hybrids (Supplementary Table 4), indicating considerable variation among the parents and hybrids for majority of the forage traits. The location \times hybrids interaction effect was significant for TDFY in the single cross hybrids, CP at first cut in both the single cross and top cross hybrids, and IVOMD at second cut in the single cross hybrids and at first cut in top cross hybrids, suggesting that the hybrids need to be evaluated under multiple sites and multiple years/seasons to identify stable cultivar. In line with the current results, Shinde (2011) reported high location \times hybrids interaction effects for CP and IVDMD (*in vitro* dry matter digestibility) in pearl millet.

Mean Performance of Hybrids and Heterosis for Forage Traits

The mean TDFY (15 t ha⁻¹) in the single cross hybrids was found to be higher than those of the top cross hybrids (TDFY: 11 t ha⁻¹) (Table 1), suggesting single cross hybrids offers opportunities for the increased forage productivity than the top cross hybrids. The BPH for TDFY ranged from 13 to 422% with a mean of 136% in single cross hybrids and 0 to 176% with a mean of 57% in the top cross hybrids (Table 2), indicating great potential for increasing the forage productivity in pearl millet. The results showed that 71 of 80 single cross hybrids (89%) and 25 of 50 top cross hybrids (50%) had greater than 35% TDFY than the better parent (Supplementary Tables 5, 6), indicating overdominance gene action plays a major role for TDFY under study. The results obtained for BPH of forage yield in this study were higher than the BPH of forage yield reported earlier in the single cross hybrids (Gupta et al., 2018; Ponnaiah et al., 2019) and in top cross or population hybrids (Bidingier et al., 2003; Pucher et al., 2016; Sattler et al., 2019) of pearl millet. Fifteen and 30 single cross and top cross hybrids significantly outyielded the check hybrid Nutrifeed by $\geq 15\%$ for TDFY, respectively (Supplementary Tables 7, 8).

Similarly, the forage quality trait CP ranged from 9 to 14% in single cross hybrids and 9 to 13% in top cross hybrids across cuts, which is still higher than the minimum requirement (about 7%) recommended for feed protein in microbes in the rumen of ruminants (Van Soest, 1994). The BPH for CP varied from -31 to 9% and -33 to 0.3% in single cross and top cross hybrids across cuts, respectively. Five (two and three at first and second

TABLE 1 | The mean performances for the forage related morphological and biochemical traits of both the single cross hybrids (SCHs) and top cross hybrids (TCHs) for both the cuts, evaluated summer season of 2015 at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, and Tamil Nadu Agricultural University (TNAU), Coimbatore.

Traits	Single cross										Top cross									
	Parents					Hybrids					Parents					Hybrids				
	Minimum	*FC	*SC	FC	SC	Average	Minimum	FC	SC	Maximum	Minimum	FC	SC	Average	Minimum	FC	SC	Maximum	Average	SC
Desirable traits																				
PH (cm)	NA	51.8	NA	104.7	NA	81.6	94.6	NA	125.8	152.1	93.9	70.5	177.0	93.3	131.4	81.1	144.9	100.9	212.0	139.9
GFY (t ha ⁻¹)	12.3	0.7	35.4	5.7	26.5	2.8	30.9	1.9	42.5	5.7	21.4	0.5	63.4	5.3	44.2	2.7	31.5	2.3	50.8	5.7
TGIFY (t ha ⁻¹)	13.6			38.8		28.0	34.0		47.6		23.0		66.2		46.7		33.8		55.0	
TDFY (t ha ⁻¹)	3.2			8.7		5.4	6.8		14.5		2.5		8.9		6.2		6.4		21.8	
DFY (t ha ⁻¹)	3.2	0.3	10.4	1.8	5.0	0.9	8.3	0.7	14.2	1.7	3.0	0.2	8.2	1.7	5.9	0.9	8.2	0.8	19.5	1.8
CP (%)	10.8	9.6	14.6	12.9	12.7	11.4	9.4	8.6	13.7	12.0	10.4	10.1	13.7	12.4	11.8	11.3	9.0	9.2	12.5	11.2
ME (MJ kg ⁻¹)	6.5	6.3	7.2	7.2	6.8	6.8	6.3	6.5	7.1	6.9	6.2	6.7	6.9	7.1	6.6	6.9	6.0	6.6	7.5	7.4
IVOMD (%)	47.9	46.4	52.9	52.8	49.9	49.5	45.4	47.1	50.8	49.8	46.9	48.2	50.4	51.9	48.3	50.2	43.8	48.1	53.3	53.2
Undesirable traits																				
NDF (%)	61.7	61.5	68.2	65.9	64.8	63.5	62.9	63.3	65.7	64.9	65.2	62.0	69.6	66.1	67.2	63.6	65.3	63.8	72.4	66.5
ADF (%)	28.4	28.0	32.5	32.4	30.9	30.5	27.7	29.8	30.7	31.9	31.7	29.1	35.1	33.3	33.3	31.2	30.0	29.3	38.3	33.2
ADL (%)	3.6	3.7	4.6	4.4	4.2	4.0	3.6	3.8	4.3	4.2	4.4	3.5	5.0	4.2	4.7	3.9	4.0	4.0	5.7	4.4

*FC: first cut; SC: second cut; PH: plant height; GFY: green forage yield; TGIFY: total green forage yield; DFY: dry forage yield; TDFY: total dry forage yield; CP: crude protein; NDF: neutral detergent fiber; ADF: acid detergent fiber; ADL: acid detergent lignin; ME: metabolizable energy; IVOMD: *in vitro* organic matter digestibility.

TABLE 2 | Heterosis for the forage linked morphological and biochemical traits in SCHs and TCHs evaluated in summer season of 2015 at ICRISAT, Patancheru, and TNAU, Coimbatore.

Traits	Materials	Cutting intervals	Better-parent heterosis (%)			Standard heterosis (%)		
			Minimum	Maximum	Average	Minimum	Maximum	Number of hybrids better than check
Desirable traits								
PH (cm)	‡SCHs	First cut (80)a	NA	NA	NA	−43.8	−1.5	0
		Second cut (45)b	4.0	73.2	37.3	−27.5	17.5	25
	†TCHs	First cut (50)c	2.3	90.1	35.1	−20.8	2.4	2
		Second cut (24)d	15.4	77.9	41.4	−17.1	17.7	21
GFY (t ha ^{−1})	SCHs	First cut	−8.2	74.4	32.6	−51.5	10.5	5
		Second cut	−37.0	301.9	51.5	−44.8	27.5	22
	TCHs	First cut	−46.1	2.9	−20.2	−27.3	12.7	6
		Second cut	−57.7	101.4	26.1	−62.8	14.2	7
TGFY (t ha ^{−1})	SCHs	Combined	1.8	118.4	47.9	−39.3	8.0	7
	TCHs	combined	−40.7	5.3	−18.1	−27.6	10.2	7
TDFY (t ha ^{−1})	SCHs	Combined	13.3	421.9	136.0	−61.8	35.2	9
	TCHs	combined	0.0	176.1	56.5	−48.1	75.8	10
DFY (t ha ^{−1})	SCHs	First cut	−13.1	344.3	154.2	−68.2	−0.9	0
		Second cut	−29.7	248.2	50.7	−37.7	11.2	9
	TCHs	First cut	16.9	318.7	75.5	−33.9	9.1	3
		Second cut	−55.1	109.3	19.5	−44.4	7.3	4
CP (%)	SCHs	First cut	−27.8	8.8	−14.4	−11.2	9.9	15
		Second cut	−30.7	4.5	−12.8	−14.0	8.8	12
	TCHs	First cut	−33.0	0.3	−15.2	−12.2	3.0	2
		Second cut	−23.2	−3.1	−14.7	−9.2	9.4	15
ME (MJ kg ^{−1})	SCHs	First cut	−6.1	7.7	1.7	−2	1.6	18
		Second cut	−8.5	3.9	−2.1	−4.1	2.3	24
	TCHs	First cut	−12.1	12.5	1.6	−2.5	1.1	14
		Second cut	−6.6	4.9	−0.8	−2.5	2.9	2
IVOMD (%)	SCHs	First cut	−8.2	6.7	0.1	−3.4	2.6	12
		Second cut	−10.8	3.4	−2.9	−2.9	1.6	21
	TCHs	First cut	−13.2	11.3	0.4	−1.2	0.6	19
		Second cut	−6.1	3.6	−1.8	−2.5	3.1	6
Undesirable traits								
NDF (%)	SCHs	First cut	−7.8	6.4	−0.4	−2.8	1.4	47
		Second cut	−3.9	7.3	1.2	−1.8	2.2	28
	TCHs	First cut	−4.5	7.6	0.2	−1.6	3.1	13
		Second cut	−1.7	6.3	2.0	−2.4	2.8	16
ADF (%)	SCHs	First cut	−13.2	3.7	−3.1	−4.2	2.6	26
		Second cut	−7.5	14.3	2.3	−1.0	2.4	20
	TCHs	First cut	−11.3	15.3	−1.5	−0.1	3.4	2
		Second cut	−10.6	6.3	−0.2	−1.3	2.2	20
ADL (%)	SCHs	First cut	−20.2	14.0	−1.9	−6.9	8.0	9
		Second cut	−12.0	19.1	2.4	−1.2	6.8	10
	TCHs	First cut	−17.3	18.5	−2.0	−2.4	14.2	4
		Second cut	−6.6	18.3	3.9	−1.3	7.5	3

a, b, c, and d indicates 80, 45, 50, and 24 hybrids, respectively; ‡ SCHs: single cross hybrids; † TCHs: top cross hybrids. PH: plant height; GFY: green forage yield; TGFY: total green forage yield; DFY: dry forage yield; TDFY: total dry forage yield; CP: crude protein; NDF: neutral detergent fiber; ADF: acid detergent fiber; ADL: acid detergent lignin; ME: metabolizable energy; IVOMD: in vitro organic matter digestibility.

cuts, respectively) single cross hybrids and one (at second cut) top cross hybrid exhibited numerically positive BPH for CP, but none were significant. The check hybrid Nutrifeed had CP of 11.9% and 10.9% in single cross hybrids and 11.5% and 10.5%

in top cross hybrids, at first and second cuts, respectively. Four and one single cross and top cross hybrids across cuts found to have significant positive or numerically higher/or at par with the check hybrid for CP, respectively. The forage IVOMD varied

from 45 to 54% and 44 to 53% across cuts in single cross and top cross hybrids, respectively. It has been reported that 1% increase in digestibility (IVOMD) in sorghum and pearl millet stovers could lead to increase in the animal outputs by 6% to 8% (Kristjanson and Zerbini, 1999). The BPH for IVOMD varied from -11 to 7% in single cross hybrids and -13 to 11% in top cross hybrids across cuts. Five single cross hybrids and three top cross hybrids exhibited numerically positive BPH for IVOMD across cuts, with one top cross hybrid significantly greater than BP. However, five single cross hybrids and three top cross hybrids had significant positive or positive BPH combined for TDFY and IVOMD for both the cuts (**Supplementary Tables 5, 6**). The check hybrid Nutrifeed had IVOMD of 51% and 48.6% at first cut and, 49.9% and 50.5% at second cut, for single cross and top cross hybrids, respectively. Ten single cross hybrids and three top cross hybrids across cuts outperformed the check hybrid (Nutrifed) for IVOMD values.

The single cross hybrids found to have higher mean values of desirable forage traits, such as TGFY, TDFY, and IVOMD, and low negative mean values of undesirable forage traits, such as NDF, ADF, and ADL at first cut than top cross hybrids (**Table 1**). This might be due to the involvement of inbreds as parents in single cross hybrids, as inbreds (hybrid parental lines) were bred for the improved forage traits. In addition, the inbred lines have favorable alleles fixed, whereas the alleles in the germplasm/OPVs are having intermediate frequencies (alleles that are not fixed), thereby the genes in top cross hybrids are less likely to combine for the favorable traits to the extent as in case of single cross hybrids (Miranda Filho, 1999; Riday et al., 2002).

Some hybrids from both the single cross and top cross hybrids had higher values for 2–3 forage traits, than the best check hybrid Nutrifed. For instance, one single cross hybrid (SCH38: L05 × T16) had significant positive standard heterosis (SH) for 12% TDFY, 4% CP and had superior or on par IVOMD with the check hybrid Nutrifed at first cut (**Supplementary Table 7**). Besides these, two hybrids SCH12: L02 × T14 and SCH62: L08 × T16 had significant positive or positive SH for TDFY (25% and 12%, respectively) with comparable IVOMD across cuts, and a hybrid SCH17: L03 × T11 had significant positive SH of 15% TDFY and comparable IVOMD at second cut. Similarly, one top cross hybrid (TCH42: L05 × T07) had significant positive or positive SH for 57% TDFY, 9% CP at second cut, and had superior or on par IVOMD with the check hybrid Nutrifed across cuts,

and a hybrid TCH45: L05 × T10 had SH of 37% TDFY and 3% IVOMD at second cut over the check hybrid (**Supplementary Table 8**). Additionally, two each of single cross (64% and 22% TDFY; 5% and 4% CP at first and second cuts, respectively) and top cross hybrids (57% and 18% TDFY; 10% and 5% CP at second cut) out yielded the check hybrid Nutrifed for these forage traits. The hybrids identified from single cross and top cross hybrids indicated that it is possible to breed the hybrids for high forage yield or for superior forage quality or combination of both forage yield and better quality in pearl millet.

Combining Ability and Gene Action

The mean squares due to GCA and SCA variances were found to be significant for TDFY in single cross hybrids, CP at first cut in single cross hybrids, and IVOMD at second cut in single cross hybrids, and also at first cut in top cross hybrids, indicating the importance of both the additive and non-additive effects for these traits (**Supplementary Table 4**). Highly significant interactions between location × (line × tester) for TDFY in single cross hybrids, CP at first cut in both single cross and top cross hybrids, and IVOMD at second cut in single cross hybrids, and at first cut in top cross hybrids indicated that hybrid performance was influenced by the locations. Baker predictability ratio (PR) was relatively lower to unity ($PR \leq 0.80$) for TDFY in both the single cross and top cross hybrids, CP at first cut in single cross hybrids, and for both the cuts in top cross hybrids, and IVOMD for both the cuts in single cross and top cross hybrids, indicating greater importance of the non-additive gene action for important forage traits under study (**Table 3** and **Supplementary Table 4**). These results are in agreement with the findings of Shinde (2011) for CP in single cross hybrids, and Ouendeba et al. (1996) for IVDMD in population hybrids, who reported that forage quality traits were under non-additive gene action in pearl millet. However, TGFY for both the single cross and top cross hybrids, and CP in single cross hybrids of second cut, were controlled by the additive gene action and thus can be improved through selection.

The GCA effects of the single cross and top cross parents for the forage quantity and quality traits under different cutting intervals are given in **Supplementary Table 9**. Seven parents (for TGFY and/or TDFY) in single cross hybrids and three parents (for TDFY) in top cross hybrids had significant positive GCA effects for these forage quantity traits, indicating that these parents can be used to enhance the forage yield potential in

TABLE 3 | Gene effects for important forage linked morphological and biochemical traits in single cross and top cross hybrids.

Traits	Additive				Non additive			
	Single cross hybrids		Top cross hybrids		Single cross hybrids		Top cross hybrids	
	First cut	Second cut	First cut	Second cut	First cut	Second cut	First cut	Second cut
TGFY (t ha ⁻¹)		√ [†]		√				
TDFY (t ha ⁻¹)						√		√
CP (%)		√			√		√	√
IVOMD (%)					√	√	√	√

√[†] indicates presence of gene action. TGFY-total green forage yield, TDFY-total dry forage yield, CP-crude protein and IVOMD-in vitro organic matter digestibility.

TABLE 4 | Associations among the hybrid performance, better-parent heterosis (BPH), and combining ability in single cross and top cross pearl millet hybrids for the forage linked traits.

Traits	F ₁ vs. GCAsum				F ₁ vs. SCA				GCA vs. BPH				SCA vs. BPH			
	Single cross hybrids		Top cross hybrids		Single cross hybrids		Top cross hybrids		Single cross hybrids		Top cross hybrids		Single cross hybrids		Top cross hybrids	
	First cut	Second cut	First cut	Second cut	First cut	Second cut	First cut	Second cut	First cut	Second cut	First cut	Second cut	First cut	Second cut	First cut	Second cut
PH (cm)	NA	0.78***	0.60***	0.64***	NA	0.62***	0.80***	0.77***	NA	0.09	0.10	0.61**	NA	0.58***	0.42***	0.65***
GFY (t ha ⁻¹)	0.82***	0.55***	0.62***	0.75***	0.57***	0.83***	0.79***	0.66***	0.43***	-0.41**	0.09	0.51*	0.53***	0.56***	0.67***	0.51*
TGFY (t ha ⁻¹)	0.63***	0.62***	0.63***	0.62***	0.71***	0.71***	0.78***	0.78***	0.10	0.15	0.35*	0.15	0.67***	0.67***	0.70***	0.70***
DFY (t ha ⁻¹)	0.48***	0.72***	0.56***	0.65***	0.88***	0.70***	0.83***	0.76***	0.30**	-0.33*	0.35*	-0.10	0.63***	0.63***	0.52***	0.60**
TDFY (t ha ⁻¹)	0.33**	0.61***	0.33**	0.61***	0.95***	0.95***	0.78***	0.78***	0.00	0.00	0.38**	0.38**	0.71***	0.71***	0.72***	0.72***
CP (%)	0.60***	0.73***	0.53***	0.69***	0.80***	0.69***	0.85***	0.73***	-0.12	0.33*	0.26	0.42*	0.67***	0.57***	0.66***	0.58***
NDF (%)	0.55***	0.57***	0.56***	0.55**	0.84***	0.82***	0.82***	0.84**	0.09	0.36*	0.45**	0.23	0.50***	0.64***	0.80***	0.55**
ADF (%)	0.59***	0.59***	0.63***	0.47*	0.81***	0.81***	0.78***	0.72***	-0.16	0.58***	0.55***	0.40	0.71***	0.40**	0.74***	0.60**
ADL (%)	0.59***	0.67***	0.56***	0.66***	0.80***	0.74***	0.83***	0.75***	-0.13	0.46**	0.46**	0.59**	0.69***	0.62***	0.75***	0.38
ME (MJ/kg)	0.51***	0.68***	0.65***	0.60**	0.83***	0.73***	0.76***	0.80***	-0.01	0.34*	0.67***	0.47*	0.84***	0.72***	0.64***	0.68***
IVOMD (%)	0.58***	0.67***	0.60***	0.46*	0.81***	0.74***	0.80***	0.80***	-0.08	0.35*	0.52***	0.27	0.79***	0.52***	0.72***	0.63***

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and NA: not available. F₁: hybrid performance, GCAsum: the sum of GCA for two parents, SCA: specific combining ability, BPH: better parent heterosis. PH: plant height; GFY: green forage yield; TGFY: total green forage yield; DFY: dry forage yield; TDFY: total dry forage yield; CP: crude protein; NDF: neutral detergent fiber; ADF: acid detergent fiber; ADL: acid detergent lignin; ME: metabolizable energy; IVOMD: *in vitro* organic matter digestibility.

breeding high biomass pearl millet. Furthermore, some parents had significant positive GCA effects for multiple forage linked traits: single cross parent (SCP)-L02 (TGFY and IVOMD at second cut), SCP-L06 (TDFY and IVOMD at first cut), and SCP-L10 (TGFY and CP at first cut) in single cross hybrids, and top cross parent (TCP)-T08 (TDFY and IVOMD at first cut) in top cross hybrids. These parents identified as good general combiners can be utilized in the breeding programs aiming to improve the forage traits in the terms of higher forage productivity coupled with superior forage quality under multi-cut system.

The estimates of SCA effects of the single cross and top cross hybrids for the forage quantity and biochemical quality traits are provided in **Supplementary Tables 10, 11**. Nineteen single cross hybrids and eight top cross hybrids were found to have highly significant positive SCA effects for the forage yield, and thus could be used in the development of high yielding forage hybrids in pearl millet. Additionally, some of the experimental hybrids had significant positive SCA effects for two or more forage traits, for instance, the hybrids (SCH05: L01 \times T15 and SCH12: L02 \times T14) for forage traits (TDFY and CP at first cut, and TDFY and IVOMD at second cut, respectively) along with high mean values for these traits. Similarly, in top cross hybrids, the hybrids (TCH18: L02 \times T13 and TCH23: L03 \times T08) at first cut and one hybrid (TCH45: L05 \times T10) at second cut had significant positive SCA effects with high mean values for forage traits (TDFY and IVOMD). The identified single cross and top cross hybrids in the present study can greatly contribute to the forage productivity in the terms of forage quantity (TDFY) and quality (CP and IVOMD).

Correlation Between Heterosis and Combining Ability of Forage Traits

The SCA effects showed significant positive correlations with hybrid performance and BPH for most of the forage traits across cuts in the single cross and top cross hybrids (**Table 4**). The GCA effects also showed positive significant correlations with hybrid performance for most of the forage traits, but low or even no

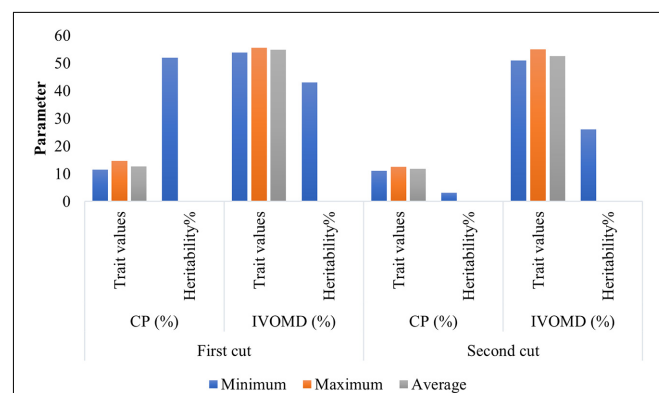


FIGURE 1 | Phenotypic performance of hybrid parents for the forage quality traits in 105 hybrid parents (Set-III) during summer seasons of 2015 and 2016 at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru.

correlations with BPH (Table 4). These results indicated that both the additive and non-additive gene effects were important for hybrid performance, while the non-additive gene effects were the major cause for heterosis. Similar findings were also reported earlier in other crops, such as in maize (Yu et al., 2020), barley (Bernhard et al., 2017) and in rice (Huang et al., 2015).

Better parent heterosis (BPH) of forage quantity (TGFY and TDFY) traits did not show correlation with the forage quality (CP and IVOMD) traits across cuts in the single cross and top cross hybrids (Supplementary Table 12), indicating that the forage quantity and quality traits can be improved independent of each other. These results are in agreement with the previous studies in pearl millet (Govintharaj et al., 2018) and in other crop, such as in sorghum (Aruna et al., 2015) which reported no associations between the forage quantity and quality traits.

Genome Wide Associations for the Forage Quality Traits

The minimum, maximum, and average values of the two traits, CP and IVOMD, observed at the different cutting intervals over 2 years are shown in Figure 1. The mean CP at first cut was almost on par with the second cut, whereas IVOMD differed by two units between the first and second cuts. The results showed huge variations between the parental lines for CP (11.5%–14.6% at first cut and 11%–12.5% at second cut) and IVOMD (53.8%–55.6% and 51%–55% at first and second cuts, respectively) as compared with the earlier studies in pearl millet (Bidinger et al., 2010; Blümmel et al., 2010). Low to moderate broad sense heritability was observed for CP and IVOMD across the two cuts (Figure 1). However, several other workers (Hash et al., 2006; Bidinger and Blümmel, 2007; Rai et al., 2012; Govintharaj et al., 2017) earlier reported low heritability for the forage quality traits in pearl millet.

Genome wide association study among 34,691 SNPs and each of the two traits, CP and IVOMD, at two different cutting intervals, detected 10 SNPs which were above the chosen

threshold level shown in the Manhattan plots (Figure 2 and Table 5), and the corresponding Q-Q plots are provided in the (Supplementary Figure 1). At second cut, during the summer season of 2016, a SNP (S4_69014036) was located on the linkage group 4 (LG4) which was found to be significantly associated with CP with the p value of $5.63\text{E-}07$. The same SNP was found to be tightly associated with CP with the p value of $9.05\text{E-}07$ when analyzed using the pooled hybrid parental data from summer seasons of 2015 and 2016. Six SNPs, one (S6_227902580) on LG6, two (S3_75463586 and S3_291119752) on LG3, and one each on LG4 (S4_46289498), LG1 (S1_180007567) and LG5 (S5_154075820), were found to be significantly associated with IVOMD at first cut during the summer season of 2015, having the p values of $1.57\text{E-}17$, $4.71\text{E-}14$, $2.53\text{E-}11$, $1.16\text{E-}10$, $1.78\text{E-}10$, and $5.59\text{E-}07$, respectively. Also, at second cut, during summer season of 2016, three SNPs S7_104645663 (LG7), S6_192886095 (LG6), and S4_51491754 (LG4) with the p values of $1.06\text{E-}09$, $4.16\text{E-}08$, and $2.21\text{E-}07$, respectively, were tightly linked with IVOMD. Similar to this study, Nepolean et al. (2006) found SSRs markers linked to the traits CP on LG4, and IVOMD on LG1, and LG6 using the bi-parental mapping population with drought tolerance background in pearl millet. No significant SNPs were found which are common for both CP and IVOMD, in both the cuts, suggesting that the genetic basis of these traits may not be the same (at different cutting intervals). These identified SNPs for the forage quality traits should be validated and then can be introgressed into the genetic background of elite/locally well adapted popular varieties to improve the forage quality in pearl millet through the marker-assisted selection.

Gene Annotation for the Forage Quality Traits

One SNP was found closely associated with CP and nine SNPs were found associated with IVOMD (Table 5). The SNP for CP corresponded to the gene Pgl_GLEAN_10029543

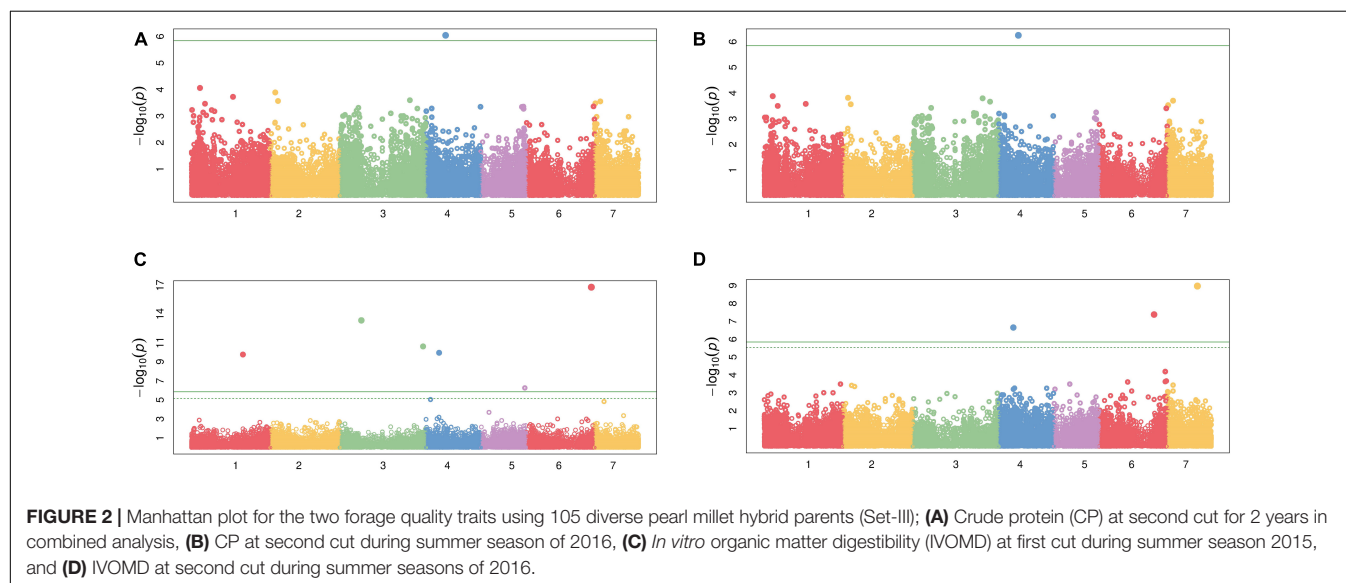


TABLE 5 | The traits linked markers, genes, and their functions identified for the forage quality traits in 105 hybrid parents evaluated during the summer seasons of 2015 and 2016 at ICRISAT, Patancheru.

S.No.	GWAS						Gene annotation			
	SNP	Chr.	Pos.	P.value	Minor allele frequency	FDR_Adjusted_P.values	Gene	Annotation	Gene function	Gene function retrieved from Genome server
CP (crude protein) at second cut in combined analysis										
1	S4_69014036	4	69014036	9.05E-07	0.242857	0.03141	Pgl_GLEAN_10029543	upstream_gene_variant	Mre11 DNA-binding presumed domain	Monocot Plaza and Phytozome
CP at second cut during summer season of 2016										
1	S4_69014036	4	69014036	5.63E-07	0.242857	0.01953	Pgl_GLEAN_10029543	upstream_gene_variant	Mre11 DNA-binding presumed domain	Monocot Plaza and Phytozome
In vitro organic matter digestibility (IVOMD) at first cut during summer season of 2015										
1	S6_227902580	6	227902580	1.57E-17	0.147619	5.46E-13	-	intergenic_region	NA	Monocot Plaza and Phytozome
2	S3_75463586	3	75463586	4.71E-14	0.138095	8.17E-10	-	intergenic_region	NA	Monocot Plaza and Phytozome
3	S3_291119752	3	291119752	2.53E-11	0.128571	2.92E-07	Pgl_GLEAN_10024973	synonymous_variant	Coatomer epsilon subunit	Monocot Plaza and Phytozome
4	S4_46289498	4	46289498	1.16E-10	0.095238	1.01E-06	Pgl_GLEAN_10027418	upstream_gene_variant	NHL domain-containing protein	Monocot Plaza and Phytozome
5	S1_180007567	1	180007567	1.78E-10	0.190476	1.24E-06	-	intergenic_region	NA	Monocot Plaza and Phytozome
6	S5_154075820	5	154075820	5.59E-07	0.114286	0.00323	Pgl_GLEAN_10037615	upstream_gene_variant	CCT motif	
IVOMD at second cut during summer season of 2016										
1	S7_104645663	7	104645663	1.06E-09	0.390476	3.69E-05	Pgl_GLEAN_10012611	intron_variant	Methyltransferase domain/Hen1 La-motif C-terminal domain	Monocot Plaza and Phytozome
2	S6_192886095	6	192886095	4.16E-08	0.114286	0.00072	Pgl_GLEAN_10034437	synonymous_variant	Coatomer beta C-terminal region	Monocot Plaza and Phytozome
3	S4_51491754	4	51491754	2.21E-07	0.414286	0.00256	Pgl_GLEAN_10032186	splice_region_variant and intron_variant	NA	Monocot Plaza and Phytozome

Chr.: chromosome; Pos.: position; NA: not available.

that coded for Mre11 DNA-binding presumed domain. Similarly, out of nine SNPs found associated with IVOMD, three were found in coding sequences (CDS) region and one in splice region which were uncharacterized for any of the gene function. The genes Pgl_GLEAN_10034437, Pgl_GLEAN_10012611, Pgl_GLEAN_10027418, Pgl_GLEAN_10024973, and Pgl_GLEAN_10037615 were found coding for functions, such as Coatomer beta C-terminal region, Methyltransferase domain/Hen1 La-motif C-terminal domain, NHL domain containing protein, Coatomer epsilon subunit, and CCT motif, respectively. Furthermore, some of the genes, such as Pgl_GLEAN_10029543, Pgl_GLEAN_10034437, Pgl_GLEAN_10012611, Pgl_GLEAN_10027418, Pgl_GLEAN_10024973, Pgl_GLEAN_10037615, and Pgl_GLEAN_10032186 were also responsible for cold tolerance (Meng et al., 2021).

CONCLUSION

Both the single cross and top cross hybrids had wide variability for the forage linked traits across cuts. The single cross hybrids had higher forage yielding traits (TGFY and TDFY), and higher IVOMD (desirable forage quality trait) and lower NDF, ADF, and ADL (undesirable forage quality traits) at first cut than top cross hybrids. The mean BPH was higher for TGFY, TDFY, and CP across cuts in single cross hybrids than top cross hybrids. Some single and top cross hybrids outperformed the commercial check hybrid for the forage yield and quality traits, these identified hybrids can be further evaluated in the multi-location trials to confirm their yield potential and stability prior to commercial release. The present study has identified the potential lines and testers for GCA effects, combined for both the forage yield and quality traits, which can offer opportunities for developing hybrids with increased forage productivity in pearl millet. Most of the forage traits across cuts in both the single cross and top cross hybrids were predominantly controlled by the non-additive gene action. No significant correlation was observed between the forage quantity and quality traits indicating that these traits can be improved independently. GWAS identified ten genomic regions associated with the forage quality traits (CP and IVOMD), and thus can be further validated, for improving the pearl millet forage quality traits through marker assisted selection. Significant genomic loci and candidate genes identified from this study lay a foundation for the development of high biomass cultivars with superior forage quality trait in the future forage breeding programs.

REFERENCES

- Aghaalkhani, M., Ahmadi, M. E., and Modarres Sanavy, A. M. (2008). Forage yield and quality of pearl millet (*Pennisetum americanum*) as influenced by plant density and nitrogen rate. *Pajouhesh Sazandegi* 77, 19–27.
- Amarendra Reddy, A., Yadav, O. P., Malik, D. P., Singh, I. P., Ardeshta, N. J., Kundu, K. K., et al. (2013). *Utilization Pattern, Demand and Supply of Pearl Millet Grain and Fodder in Western India*. Working Paper Series no. 37. Patancheru: International Crops Research Institute for the Semi-Arid Tropics, 24.
- Amer, S., and Mustafa, A. F. (2010). Effects of feeding pearl millet silage on milk production of lactating dairy cows. *J. Dairy Sci.* 93, 5921–5925. doi: 10.3168/jds.2010-3279
- Aruna, C., Swarnalatha, M., Kumar, P. P., Devender, V., Suguna, M., Blümmel, M., et al. (2015). Genetic options for improving fodder yield and quality in forage sorghum. *Tropical Grasslands – Forrajes Tropicales* 3, 49–58. doi: 10.17138/TGFT(3)49-58
- Arunachalam, V. (1974). The fallacy behind the use of a modified Line × Tester design. *Indian J. Genet. Plant Breed.* 34, 280–287.
- Arya, R. K., Kumar, S., Yadav, A. K., and Kumar, A. (2013). Grain quality improvement in pearl millet: a review. *Forage Res.* 38, 189–201.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <http://dataverse.icrisat.org/dataset.xhtml?persistentId=doi:10.21421/D2/IWXCUIJ>.

AUTHOR CONTRIBUTIONS

SG, PG, MM, PS, and MB designed the experiments. AV, SK, AR, SS, and RV helped in data analysis. RV involved in genotyping for forage type hybrid parents. MB involved in phenotyping for forage quality traits by NIRS. PG and SK wrote the manuscript. MM and PS reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

Financial support of the project was provided by the ICRISAT-Pearl Millet Hybrid Parents Research Consortium (PMHPRC) and conducted under the CGIAR, Research Program on Grain Legumes and Dryland Cereals (CRP-GLDC) to the PG to carry out this research work as a part of doctoral program.

ACKNOWLEDGMENTS

This research work was carried out as part of a Ph.D. thesis of PG, submitted to Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India. It was financially supported by the ICRISAT-Pearl Millet Hybrid Parents Research Consortium (PMHPRC) and conducted under the CGIAR, Research Program on Grain Legumes and Dryland Cereals (CRP-GLDC). We also acknowledge K.V.S. Prasad and Ramakrishna Reddy from ICRISAT-ILRI for sample preparation and NIRS analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.687859/full#supplementary-material>

- Babiker, S. A., Khair, M. A. M., and Tahir, I. S. A. (2014). Exploitation of forage attribute-based variations in Sudan pearl millet [*Pennisetum glaucum* (L.) R. Br.] collections. *Plant Genet. Resour.* 12, 83–90. doi: 10.1017/S1479262113000312
- Baker, R. J. (1978). Issues in diallel analysis. *Crop Sci.* 18, 533–536. doi: 10.2135/cropsci1978.0011183X001800040001x
- Baltensperger, D. D. (2002). “Progress with proso, pearl and other millets,” in *Trends in New Crops and New Uses*, eds J. Janick and A. Whipkey (Alexandria, VA: ASHS Press), 100–103.
- Bernhard, T., Friedt, W., Voss-Fels, K. P., Frisch, M., Snowden, R. J., and Wittkop, B. (2017). Heterosis for biomass and grain yield facilitates breeding of productive dual-purpose winter barley hybrids. *Crop Sci.* 57, 2405–2418. doi: 10.2135/cropsci2016.10.0872
- Bidinger, F. R., and Blümmel, M. (2007). Determinants of ruminant nutritional quality of pearl millet [*Pennisetum glaucum* (L.) R. Br.] stover. I. Effects of management alternatives on stover quality and productivity. *Field Crops Res.* 103, 119–128. doi: 10.1016/j.fcr.2007.05.006
- Bidinger, F. R., Blümmel, M., Hash, C. T., and Choudhary, S. (2010). Genetic enhancement for superior food-feed traits in a Pearl Millet [*Pennisetum glaucum* (L.) R. Br.] variety by recurrent selection. *Anim. Nutr. Feed Techn.* 10S, 61–68.
- Bidinger, F. R., Yadav, O. P., Sharma, M. M., van Oosterom, E. J., and Yadav, Y. P. (2003). Exploitation of heterosis for simultaneous improvement in both grain and stover yields of arid zone pearl millet (*Pennisetum glaucum* (L.) R. Br.). *Field Crops Res.* 83, 13–26. doi: 10.1016/S0378-4290(03)00006-6
- Blümmel, M., Bidinger, F. R., and Hash, C. T. (2007). Management and cultivar effect on ruminant nutritional quality of pearl millet [*Pennisetum glaucum* (L.) R. Br.] stover. II. Effects of cultivar choice on stover quality and productivity. *Field Crops Res.* 103, 129–138. doi: 10.1016/j.fcr.2007.05.007
- Blümmel, M., Khan, A. A., Vadez, V., Hash, C. T., and Rai, K. N. (2010). Variability in stover quality traits in commercial hybrids of pearl millet (*Pennisetum glaucum* (L.) R. Br.) and grain stover trait relationships. *Anim. Nutr. Feed Techn.* 10S, 29–38.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Brunette, T., Baurhoo, B., and Mustafa, A. F. (2014). Replacing corn silage with different forage millet cultivars: effects on milk yield, nutrient digestion, and ruminal fermentation of lactating dairy cows. *J. Dairy Sci.* 97, 6440–6449. doi: 10.3168/jds.2014-7998
- Burton, G. W. (1965). Pearl millet Tift 23A released. *Crops Soils* 17:19.
- Cingolani, P., Platts, A., Wang Le, L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6, 80–92.
- Dias-Martins, A. M., Pessanha, K. L., Pacheco, S., Rodrigues, J. A. S., and Carvalho, C. W. P. (2018). Potential use of pearl millet (*Pennisetum glaucum* (L.) R. Br.) in Brazil: food security, processing, health benefits and nutritional products. *Food Res. Int.* 109, 175–186. doi: 10.1016/j.foodres.2018.04.023
- Dutta, S., Sattler, F. T., Pucher, A., Inoussa, D., Issaka, A., Sy, O., et al. (2021). *Heterosis and Combining Abilities in a Diverse Seven-Parent Pearl Millet Population Diallel Tested in West Africa*. Durham, NC: Research Square.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379
- Food and Agriculture Organization [FAO] (2002). *World Agriculture: Towards 2015/2030 (Summary Report)*. Available online at: <http://www.fao.org/global-perspectives-studies/resources/detail/en/c/411230/> (accessed September 15, 2021).
- Govintharaj, P., Gupta, S. K., Blümmel, M., Maheswaran, M., Sumathi, P., Atkari, D. G., et al. (2018). Genotypic variation in forage linked morphological and biochemical traits in hybrid parents of pearl millet. *Anim. Nutr. Feed Techn.* 18, 163–175. doi: 10.5958/0974-181X.2018.00016.1
- Govintharaj, P., Gupta, S. K., Maheswaran, M., and Sumathi, P. (2017). Genetic variability, heritability and genetic advance studies in forage type hybrid parents of pearl millet. *Electron. J. Plant Breed.* 8, 1265–1274. doi: 10.5958/0975-928X.2017.00181.8
- Gupta, S. K., Ghouse, S. K. C., Atkari, D. G., and Blümmel, M. (2015). “Pearl millet with higher stover yield and better forage quality: identification of new germplasm and cultivars,” in *Proceeding of the 3rd Conference of Cereal Biotechnology and Breeding/CBB3*. November 2–4, Berlin, Germany, (Berlin).
- Gupta, S. K., Nepolean, T., Shaikh, C. G., Rai, K. N., Hash, C. T., Das, R. R., et al. (2018). Phenotypic and molecular diversity-based prediction of heterosis in pearl millet (*Pennisetum glaucum* L. (R.) Br.). *Crop J.* 6, 271–281. doi: 10.1016/j.cj.2017.09.008
- Gupta, S. K., Patil, K. S., Rathore, A., Yadav, D. V., Sharma, L. D., Mungra, K. D., et al. (2020). Identification of heterotic groups in South-Asian bred hybrid parents of pearl millet. *Theor. Appl. Genet.* 133, 873–888. doi: 10.1007/s00122-019-03512-z
- Hanna, W. (1996). “Improvement of millets: emerging trends,” in *Proceedings of the 2nd International Crop Science Congress*, New Delhi, India, eds V. L. Chopra, R. B. Singh, and A. Varma (New Delhi: Oxford & IBH Publishing Co. Pvt. Ltd.), 139–146.
- Hash, C. T., Bhaskar Raj, A. G., Lindup, S., Sharma, A., Beniwal, C. R., Folkertsma, R. T., et al. (2003). Opportunities for marker-assisted selection (MAS) to improve the feed quality of crop residues in pearl millet and sorghum. *Field Crops Res.* 84, 79–88. doi: 10.1016/S0378-4290(03)00142-4
- Hash, C. T., Blümmel, M., and Bidinger, F. R. (2006). Genotype × Environment interactions in food-feed traits in pearl millet cultivars. *Int. Sorghum Millets Newsl.* 47, 153–157.
- Huang, M., Chen, L. Y., and Chen, Z. Q. (2015). Diallel analysis of combining ability and heterosis for yield and yield components in rice by using positive loci. *Euphytica* 205, 37–50. doi: 10.1007/s10681-015-1381-8
- Kelley, T. G., Parthasarathy Rao, P., Weltzien, R. E., and Purohit, M. L. (1996). Adoption of improved cultivars of pearl millet in an arid environment, straw quality and quality considerations in western Rajasthan. *Exp. Agric.* 32, 161–171.
- Kempthorne, O. (1957). *An Introduction to Genetic Statistics*. New York, NY: John Wiley and Sons.
- Kristjansson, P. M., and Zerbini, E. (1999). *Genetic Enhancement of Sorghum and Millet Residues Fed to Ruminants: An ex Ante Assessment of Returns to Research*. Impact Assessment Series No. 3. Kenya: International Livestock Research Institute (ILRI) (aka ILCA and ILRAD), 44.
- Lauriault, L. M., Leah, H. S., Shad, H. C., and Eric, J. S. (2021). A comparison of pearl millet and sorghum–sudangrass pastures during the frost-prone autumn for growing beef cattle in semiarid region. *Agriculture* 11:541. doi: 10.3390/agriculture11060541
- Li, B., Ishii, Y., Idota, S., Tobisa, M., Niimi, M., Yang, Y., et al. (2019). Yield and quality of forages in a triple cropping system in Southern Kyushu, Japan. *Agronomy* 9:277. doi: 10.3390/agronomy9060277
- Meng, X., Liang, Z., Dai, X., Zhang, Y., Mahboub, S., Ngu, D. W., et al. (2021). Predicting transcriptional responses to cold stress across plant species. *Proc. Natl. Acad. Sci. U. S. A.* 9:e2026330118. doi: 10.1073/pnas.2026330118
- Miranda Filho, J. B. (1999). “Inbreeding depression and heterosis,” in *Genetic and Exploitation of Heterosis in Crops*, eds J. G. Coors and S. Pandey (Madison: American Society Agronomy), 69–80.
- Nepolean, T., Blümmel, M., Bhaskar Raj, A. G., Rajaram, V., Senthilvel, S., and Hash, C. T. (2006). QTLs controlling yield and stover quality traits in pearl millet. *J. Sat Agric. Res.* 2, 1–4.
- Newman, Y., Jennings, E., Vendramini, J., and Blount, A. (2010). *Pearl Millet (Pennisetum glaucum): Overview and Management*. SS-AGR-337, One of a Series of the Agronomy Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida. Florida, FL: University of Florida.
- Nurbekov, A., Jamoliddinov, A., Joldoshev, K., Rischkovsk, B., Nishanov, N., Rai, K. N., et al. (2013). Potential of pearl millet as a forage crop in wheat-based double cropping system in Central Asia. *J. SAT Agric. Res.* 11, 1–5.
- Ouendeba, B., Hanna, W. W., Ejeta, G., Nyquist, W. E., and Santini, J. (1996). Forage yield and digestibility of African pearl millet landraces in diallel with missing cross. *Crop Sci.* 36, 1517–1520. doi: 10.2135/cropsci1996.0011183X0036000600018x
- Patil, K. S., Gupta, S. K., Marathi, B., Danam, S., Thatikunta, R., Rathore, A., et al. (2020). African and Asian origin pearl millet populations: genetic diversity pattern and its association with yield heterosis. *Crop Sci.* 60, 3035–3048. doi: 10.1002/csc2.20245

- Ponnaiah, G., Gupta, S. K., Blümmel, M., Marappa, M., Pichaikannu, S., Das, R. R., et al. (2019). Utilization of molecular marker based genetic diversity patterns in hybrid parents to develop better forage quality multi-cut hybrids in pearl millet. *Agriculture* 9:97. doi: 10.3390/agriculture9050097
- Pucher, A., Sy, O., Sanogo, M. D., Angarawai, I. I., Zangre, R., Ouedrago, M., et al. (2016). Combining ability patterns among West African pearl millet landraces and prospects for pearl millet hybrid breeding. *Field Crops Res.* 195, 9–20. doi: 10.1016/j.fcr.2016.04.035
- Rai, K. N., Blümmel, M., Singh, A. K., and Rao, A. S. (2012). Variability and relationships among forage yield and quality traits in pearl millet. *Eur. J. Plant Sci. Biotechnol.* 6, 118–124.
- Rai, K. N., Reddy, B. V. S., Saxena, K. B., and Gowda, C. L. L. (2004). “Prospects of breeding sorghum, pearl millet and pigeon pea for high forage yield and quality,” in *Proceedings of the 4th International Crop Science Congress*, (Brisbane).
- Riday, H., Brummer, E. C., and Moore, K. J. (2002). Heterosis of forage quality in alfalfa. *Crop Sci.* 42, 1088–1093. doi: 10.2135/cropsci2002.1088
- SAS Institute Inc. (2017). *SAS/STAT® 14.1 User's Guide*. Cary, NC: SAS Institute Inc.
- Sattler, F. T., and Haussmann, B. I. G. (2020). A unified strategy for West African pearl millet hybrid and heterotic group development. *Crop Sci.* 60, 1–13. doi: 10.1002/csc2.20033
- Sattler, F. T., Pucher, A., Ango, I. K., Sy, O., Ahmadou, I., Hash, C. T., et al. (2019). Identification of combining ability patterns for pearl millet hybrid breeding in West Africa. *Crop Sci.* 59, 1590–1603. doi: 10.2135/cropsci2018.12.0727
- Segura, V., Vilhjalmsón, B. J., Platt, A., Korte, A., Seren, U., Long, Q., et al. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat. Genet.* 44, 825–830. doi: 10.1038/ng.2314
- Sheahan, C. M. (2014). *Plant Guide for Pearl Millet (Pennisetum glaucum)*. Cape May, NJ: USDA-Natural Resources Conservation Service.
- Shinde, G. C. (2011). *Genetic Studies and Molecular Analysis for Forage Yield in Pearl Millet (Pennisetum glaucum, L.)* R. Br. Ph.D. thesis. Rahuri: Mahatma Phule Krishi Vidyapeeth.
- Singh, S., and Gupta, S. K. (2019). Formation of heterotic pools and understanding relationship between molecular divergence and heterosis in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *PLoS One* 14:e0207463. doi: 10.1371/journal.pone.0207463
- Urrutia, J. M., José Hernández, A. A., Rosales, C. A. N., and Cervantes, B. J. F. (2015). Forage production and nutritional content of silage from three varieties of pearl millet (*Pennisetum glaucum*) harvested at two maturity stages. *J. Anim. Plant Sci.* 27, 4161–4169.
- Vadez, V., Hash, T., Bidinger, F. R., and Kholova, J. (2012). Phenotyping pearl millet for adaptation to drought. *Front. Physiol.* 3:386. doi: 10.3389/fphys.2012.00386
- Van Soest, P. J. (1994). *Nutritional Ecology of Ruminant*, 2nd Edn. Ithaca, NY: Cornell University Press.
- Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976. doi: 10.1038/nbt.3943
- Vinutha, K. S., Khan, A. A., Ravi, D., Prasad, K. V. S. V., Ramana Reddy, Y., Jones, C. S., et al. (2021). Comparative evaluation of sorghum and pearl millet forage silages with maize. *Anim. Nutr. Feed Techn.* 21, 1–14.
- Wang, J., and Zhang, Z. (2020). GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *bioRxiv [preprint]* 2020.11.29.403170. doi: 10.1101/2020.11.29.403170
- Yadav, R. S., Bidinger, F. R., Hash, C. T., Yadav, Y. P., Yadav, O. P., Bhatnagar, S. K., et al. (2002). Mapping and characterization of QTL × E interactions for traits determining grain and stover yield in pearl millet. *Theor. Appl. Genet.* 106, 512–520. doi: 10.1007/s00122-002-1081-3
- Yu, K., Wang, H., Liu, X., Xu, C., Li, Z., Xu, X., et al. (2020). Large-scale analysis of combining ability and heterosis for development of hybrid maize breeding strategies using diverse germplasm resources. *Front. Plant Sci.* 11:660. doi: 10.3389/fpls.2020.00660

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past collaboration with one of the authors, RV.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Govintharaj, Maheswaran, Blümmel, Sumathi, Vemula, Rathore, Sivasubramani, Kale, Varshney and Gupta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: frontiersin.org/about/contact



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership