

Advances in breeding for quantitative disease resistance

Edited by

Valerio Hoyos-Villegas, Harsh Raman, Anna Maria Mastrangelo and Jianjun Chen

Published in

Frontiers in Plant Science



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 978-2-88976-091-6
DOI 10.3389/978-2-88976-091-6

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

Advances in breeding for quantitative disease resistance

Topic editors

Valerio Hoyos-Villegas — McGill University, Canada

Harsh Raman — Department of Primary Industries, NSW Government, Australia

Anna Maria Mastrangelo — Research Centre for Cereal and Industrial Crops,
Council for Agricultural and Economics Research (CREA), Italy

Jianjun Chen — University of Florida, United States

Citation

Hoyos-Villegas, V., Raman, H., Mastrangelo, A. M., Chen, J., eds. (2023). *Advances in breeding for quantitative disease resistance*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-091-6

Table of contents

- 07 **Editorial: Advances in Breeding for Quantitative Disease Resistance**
Valerio Hoyos-Villegas, Jianjun Chen, Anna Maria Mastrangelo and Harsh Raman
- 12 **Characterization of the Genetic Architecture for Fusarium Head Blight Resistance in Durum Wheat: The Complex Association of Resistance, Flowering Time, and Height Genes**
Yuefeng Ruan, Wentao Zhang, Ron E. Knox, Samia Berraies, Heather L. Campbell, Raja Ragupathy, Kerry Boyle, Brittany Polley, Maria Antonia Henriquez, Andrew Burt, Santosh Kumar, Richard D. Cuthbert, Pierre R. Fobert, Hermann Buerstmayr and Ron M. DePauw
- 29 **Genomic Breeding for Diameter Growth and Tolerance to *Leptocybe* Gall Wasp and *Botryosphaeria/Teratosphaeria* Fungal Disease Complex in *Eucalyptus grandis***
Makobatjatji M. Mphahlele, Fikret Isik, Gary R. Hodge and Alexander A. Myburg
- 44 **Dissecting Quantitative Trait Loci for Spot Blotch Resistance in South Asia Using Two Wheat Recombinant Inbred Line Populations**
Chandan Roy, Navin C. Gahtyari, Xinyao He, Vinod K. Mishra, Ramesh Chand, Arun K. Joshi and Pawan K. Singh
- 53 **Genome-Wide Association Studies Reveal All-Stage Rust Resistance Loci in Elite Durum Wheat Genotypes**
Meriem Aoun, Matthew N. Rouse, James A. Kolmer, Ajay Kumar and Elias M. Elias
- 73 **Genetic Improvement for Resistance to Black Sigatoka in Bananas: A Systematic Review**
Julianna M. S. Soares, Anelita J. Rocha, Fernanda S. Nascimento, Adriadna S. Santos, Robert N. G. Miller, Cláudia F. Ferreira, Fernando Haddad, Vanusia B. O. Amorim and Edson P. Amorim
- 88 **Introgression of Maize Lethal Necrosis Resistance Quantitative Trait Loci Into Susceptible Maize Populations and Validation of the Resistance Under Field Conditions in Naivasha, Kenya**
Luka A. O. Awata, Beatrice E. Ifie, Eric Danquah, MacDonald Bright Jumbo, L. Mahabaleswara Suresh, Manje Gowda, Philip W. Marchelo-Dragga, Michael Scott Olsen, Oluwaseyi Shorinola, Nasser Kouadio Yao, Prasanna M. Boddupalli and Pangirayi B. Tongoona
- 102 **Novel Genomic Regions of *Fusarium* Wilt Resistance in Bottle Gourd [*Lagenaria siceraria* (Mol.) Standl.] Discovered in Genome-Wide Association Study**
Yanwei Li, Ying Wang, Xinyi Wu, Jian Wang, Xiaohua Wu, Baogen Wang, Zhongfu Lu and Guojing Li

- 114 **Characterization and Mapping of Spot Blotch in *Triticum durum*–*Aegilops speltoides* Introgression Lines Using SNP Markers**
Jashanpreet Kaur, Jaspal Kaur, Guriqbal Singh Dhillon, Harmandeep Kaur, Jasvir Singh, Ritu Bala, Puja Srivastava, Satinder Kaur, Achla Sharma and Parveen Chhuneja
- 127 **Genome-Wide Association Study and Genomic Prediction for Soybean Cyst Nematode Resistance in USDA Common Bean (*Phaseolus vulgaris*) Core Collection**
Ainong Shi, Paul Gepts, Qijian Song, Haizheng Xiong, Thomas E. Michaels and Senyu Chen
- 153 **Genetic Gains for Yield and Virus Disease Resistance of Cassava Varieties Developed Over the Last Eight Decades in Uganda**
Francis Manze, Patrick Rubaihayo, Alfred Ozimati, Paul Gibson, Williams Esuma, Anton Bua, Titus Alicai, Chris Omongo and Robert S. Kawuki
- 164 **Fine Mapping of the Leaf Rust Resistance Gene *Lr65* in Spelt Wheat 'Altgold'**
Qiang Zhang, Wenxin Wei, Xiangxi Zuansun, Shengnan Zhang, Chen Wang, Nannan Liu, Lina Qiu, Weidong Wang, Weilong Guo, Jun Ma, Huiru Peng, Zhaorong Hu, Qixin Sun and Chaojie Xie
- 174 **Outlook of Cassava Brown Streak Disease Assessment: Perspectives of the Screening Methods of Breeders and Pathologists**
Alfred A. Ozimati, Williams Esuma, Titus Alicai, Jean-Luc Jannink, Chiedozie Egesi and Robert Kawuki
- 182 **Pre-emptive Breeding Against Karnal Bunt Infection in Common Wheat: Combining Genomic and Agronomic Information to Identify Suitable Parents**
Livinus Emebiri, Shane Hildebrand, Mui-Keng Tan, Philomin Juliana, Pawan K. Singh, Guillermo Fuentes-Davila and Ravi P. Singh
- 193 **Breeding With Major and Minor Genes: Genomic Selection for Quantitative Disease Resistance**
Lance F. Merrick, Adrienne B. Burke, Xianming Chen and Arron H. Carter
- 215 **Genomic-Assisted Marker Development Suitable for *CsCvy-1* Selection in Cucumber Breeding**
Erdem Kahveci, Zübeyir Devran, Ercan Özkaynak, Yiguo Hong, David J. Studholme and Mahmut Tör
- 223 **Genome-Wide Association Study Reveals Novel Genetic Loci for Quantitative Resistance to Septoria Tritici Blotch in Wheat (*Triticum aestivum* L.)**
Tilahun Mekonnen, Clay H. Sneller, Teklehaimanot Haileselassie, Cathrine Ziyomo, Bekele G. Abeyo, Stephen B. Goodwin, Dagnachew Lule and Kassahun Tesfaye

- 243 **Associated SNPs, Heritabilities, Trait Correlations, and Genomic Breeding Values for Resistance in Snap Beans (*Phaseolus vulgaris* L.) to Root Rot Caused by *Fusarium solani* (Mart.) f. sp. *phaseoli* (Burkholder)**
Abigail R. Huster, Lyle T. Wallace and James R. Myers
- 260 **Identification and Validation of Genomic Regions Associated With Charcoal Rot Resistance in Tropical Maize by Genome-Wide Association and Linkage Mapping**
Zerka Rashid, Harleen Kaur, Veerendra Babu, Pradeep Kumar Singh, Sharanappa I. Harlapur and Sudha K. Nair
- 273 **Molecular Mapping and Analysis of an Excellent Quantitative Trait Loci Conferring Adult-Plant Resistance to Stripe Rust in Chinese Wheat Landrace Gaoxianguangtoumai**
Yuqi Wang, Fengying Liang, Fangnian Guan, Fangjie Yao, Li Long, Xuyang Zhao, Luyao Duan, Yu Wu, Hao Li, Wei Li, Qiantao Jiang, Yuming Wei, Jian Ma, Pengfei Qi, Mei Deng, Youliang Zheng, Houyang Kang, Yunfeng Jiang and Guoyue Chen
- 285 **Predicting Fusarium Head Blight Resistance for Advanced Trials in a Soft Red Winter Wheat Breeding Program With Genomic Selection**
Dylan L. Larkin, Richard Esten Mason, David E. Moon, Amanda L. Holder, Brian P. Ward and Gina Brown-Guedira
- 300 **Investigation and Genome-Wide Association Analysis of Fusarium Seedling Blight Resistance in Chinese Elite Wheat Lines**
Yike Liu, Guang Zhu, Zhangwang Zhu, Lin Chen, Hongli Niu, Weijie He, Hanwen Tong, Jinghan Song, Yuqing Zhang, Dongfang Ma and Chunbao Gao
- 311 **Identification of Two Major QTLs in *Brassica napus* Lines With Introgressed Clubroot Resistance From Turnip Cultivar ECD01**
Fengqun Yu, Yan Zhang, Jinghe Wang, Qilin Chen, Md. Masud Karim, Bruce D. Gossen and Gary Peng
- 323 **Development and Genetic Characterization of Peanut Advanced Backcross Lines That Incorporate Root-Knot Nematode Resistance From *Arachis stenosperma***
Carolina Ballén-Taborda, Ye Chu, Peggy Ozias-Akins, C. Corley Holbrook, Patricia Timper, Scott A. Jackson, David J. Bertoli and Soraya C. M. Leal-Bertioli
- 337 **Identification of Quantitative Trait Loci Associated With Partial Resistance to Fusarium Root Rot and Wilt Caused by *Fusarium graminearum* in Field Pea**
Longfei Wu, Rudolph Fredua-Agyeman, Stephen E. Strelkov, Kan-Fa Chang and Sheau-Fang Hwang
- 353 **Quantitative Trait Locus Mapping for Resistance Against *Pyrenopeziza brassicae* Derived From a *Brassica napus* Secondary Gene Pool**
Chinthani S. Karandeni Dewage, Katherine Cools, Henrik U. Stotz, Aiming Qi, Yong-Ju Huang, Rachel Wells and Bruce D. L. Fitt

- 366 Single Nucleotide Polymorphism Detection for Peach Gummosis Disease Resistance by Genome-Wide Association Study**
Xiongwei Li, Jiabo Wang, Mingshen Su, Jingyi Zhou, Minghao Zhang, Jihong Du, Huijuan Zhou, Kexin Gan, Jing Jin, Xianan Zhang, Ke Cao, Weichao Fang, Lirong Wang, Huijuan Jia, Zhongshan Gao and Zhengwen Ye
- 381 Influence of Elevated Temperatures on Resistance Against Phoma Stem Canker in Oilseed Rape**
Katherine Noel, Aiming Qi, Lakshmi Harika Gajula, Craig Padley, Steffen Rietz, Yong-Ju Huang, Bruce D. L. Fitt and Henrik U. Stotz



Editorial: Advances in Breeding for Quantitative Disease Resistance

Valerio Hoyos-Villegas^{1*}, Jianjun Chen², Anna Maria Mastrangelo³ and Harsh Raman⁴

¹ Department of Plant Science, McGill University, Montreal, QC, Canada, ² Mid-Florida Research & Education Center, University of Florida, Apopka, FL, United States, ³ Research Centre for Cereal and Industrial Crops, Council for Agricultural and Economics Research (CREA), Foggia, Italy, ⁴ New South Wales Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia

Keywords: quantitative disease resistance, plant breeding, genome-wide association, quantitative trait loci (QTL) analysis, pathosystem, fine mapping, pre-breeding, genomic selection (GS)

Editorial on the Research Topic

Advances in Breeding for Quantitative Disease Resistance

In plant breeding and genetics, traits are frequently classified into qualitative and quantitative. A qualitative trait is generally controlled by one or a few genes, whereas a quantitative trait is controlled by several genes. The effect of each of the alleles responsible for a quantitative trait is typically small when compared to the effect of the environment, making the inference of an individual genotype difficult to establish. Genetic bases of quantitative traits are characterized by a continuous distribution of phenotypes and detected by quantitative trait loci (QTL) analysis and genome-wide association studies (GWAS). In the world of disease resistance, quantitative disease resistance (QDR) has been reported in a large number of crops, and molecular markers tightly linked to quantitative resistance loci (QRL) controlling QDR.

Quantitative disease resistance is expressed when host plants exhibit a reduced disease reaction but not complete resistance. It is widely recognized that QDR provides long term host defense toward the disease, probably due to multiple genes requiring mutation for resistance breakdown as opposed to single genes as in the case of gene-for-gene resistance. However, QDR has been a longstanding challenge in the development of cultivars with durable resistance and new techniques such as GWAS could complement QTL mapping results. Emerging genetic, metabolomics, genomics, phenomics, machine learning, and synthetic biology tools could speed-up the development of new plant cultivars having quantitative disease resistance.

A collection of reports was assembled to represent achievements in understanding and improving QDR. New technologies provide avenues for measuring QDR in plant breeding populations, and new insights on plant-pathogen interactions provide new alternatives for studying QDR. Researchers around the world have made progress toward the goal of achieving QDR, and new tools technologies and knowledge to increase food productivity and sustainability using precision breeding to boost QDR.

The objective of this Research Topic was to collate articles updating the status of breeding for QDR. The interest was to provide an updated view of the science of breeding for QDR as well as the tools that have become available in the development of QDR. We received a total of 37 submissions, of which 27 were accepted into the collection. A group of 50 authors contributed to the collection. Among the accepted submissions, the following eight topics were covered: QTL mapping (5 articles), fine mapping (1 articles), genome-wide association (8 articles), genomic selection (4 articles), marker development (2 articles), pathogen-environment-genotype (2 articles), breeding and pre-breeding (3 articles), and reviews (2 articles).

OPEN ACCESS

Edited and reviewed by:

Salvatore Ceccarelli,
Bioversity International, Italy

*Correspondence:

Valerio Hoyos-Villegas
valerio.hoyos-villegas@mcgill.ca

Specialty section:

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

Received: 04 March 2022

Accepted: 23 March 2022

Published: 13 April 2022

Citation:

Hoyos-Villegas V, Chen J,
Mastrangelo AM and Raman H (2022)
Editorial: Advances in Breeding for
Quantitative Disease Resistance.
Front. Plant Sci. 13:890002.
doi: 10.3389/fpls.2022.890002

QTL MAPPING

This collection reported the use of novel genetic populations for the exploration of QTL in mapping, breeding and pre-breeding populations. Karandeni-Dewage et al. screened a doubled haploid (DH) population derived from the secondary gene pool of *Brassica napus* with the aim of introgressing resistance to *Pyrenopeziza brassicae*, the authors identified four QTLs that had moderate to large allelic effects. Similarly, Yu et al. studied a DH population for mapping resistance to clubroot disease, caused by the fungal pathogen, *Plasmodiophora brassicae* and used *Brassica rapa* as a source for resistance. The authors found gene-for-gene interaction with various pathotypes and identified two QTL associated with resistance.

Several QTL analysis papers in wheat diseases were published. Wu et al. studied partial resistance to five fungal isolates representing *Fusarium solani*, *F. avenaceum*, *F. acuminatum*, *F. proliferatum* and *F. graminearum* in field pea. The authors used a mapping population between resistant and susceptible parents and found multiple stable QTL for resistance while screening for the various *Fusarium* isolates. Wang et al. identified and validated stable QTL conferring adult-plant resistance to stripe rust (*Puccinia striiformis* f. sp. *tritici*), found in the Chinese landrace "Gaoxianguangtoumai." In particular, QTL *QYr.GX-2AS* was found to be present only in low frequency (5.3%) among 325 Chinese landraces. Roy et al., studied QTL for resistance to spot blotch in wheat. They used two bi-parental mapping populations and found several QTL having low to moderate effects.

FINE MAPPING

A significant contribution by Zhang et al. was published. Their paper focused on fine mapping of the leaf rust resistance gene *Lr65* in the spelt wheat cultivar "Altgold." The authors delimited *Lr65*—a 0.8 cM interval and provided one simple sequence repeat marker and a high-resolution map, further reducing the region to 60.11 Kb in size.

GENOME-WIDE ASSOCIATION ANALYSIS

Liu et al. used a set of 240 Chinese elite cultivars genotyped using a 90 K single nucleotide polymorphism (SNP) array with the aim of finding signals associated with *Fusarium* seedling blight resistance. The authors found six stable QTL accounting 4.8–7.5% of the phenotypic variation. The authors also report four Kompetitive allele specific PCR (KASP) markers to enable marker assisted selection in wheat breeding programs. Rashid et al. used a panel of 396 tropical adapted (Asian environments) maize lines genotyped with 296 k SNPs using genotyping by sequencing (GBS) approach to screen for Charcoal rot resistance (*Macrophomina phaseolina*). The authors found 19 SNPs with significant associations and developed two $F_{2:3}$ populations to validate the signals. Two QTL co-located with two of the SNP and haplotypes detected. The authors reported that many of the signals found overlap with previously reported QTL for *Gibberella* stalk rot resistance, thus

increasing the opportunity to develop resistance to multiple stalk rots.

Mekonnen et al. studied *Septoria tritici* blotch (*Zymoseptoria tritici*) in wheat and used a set of 178 bread wheat genotypes to screen for adult plant resistance and agronomic traits for 2 years. The association panel was genotyped using GBS and this resulted in 7k polymorphic SNPs. Significant marker-trait associations were found in 27% of the marker pairs, suggesting 33 putative QTL with 5 QTL reported as novel. The putative QTL explained 2.7–13.2% of total phenotypic variation.

Kaur et al. deployed 89 backcross introgression lines between *Triticum durum* × *Aegilops speltoides* and evaluated them for spot blotch resistance, caused by *Bipolaris sorokiniana* for four consecutive years. The authors identified five QTLs linked to spot blotch. In particular, QTL *Q.Sb.pau-5B* was validated in this study, serving as a future diagnostic marker for spot blotch resistance.

Li, Y., et al. used a set of 89 bottle gourd [*Lagenaria siceraria* (Mol.) Standl. accessions with the aim of finding significant associations for resistance to *Fusarium* wilt. The study genotyped the panel with 5k SNPs and revealed a total of 10 SNPs detected in at least two environments. Aoun et al. identified 64 marker trait associations (MTA) for leaf rust (*Puccinia triticina*), 46 MTAs for stripe rust (*Puccinia striiformis* f. sp. *tritici*) and 260 MTAs for stem rust (*Puccinia graminis* f. sp. *tritici*) resistance in an elite durum wheat association panel genotyped with a 90k SNP array. None of the signals for stripe rust found here corresponded to existing designations of resistance genes. In contrast, two and four of the signals for leaf rust and stem rust overlapped with known resistance genes, respectively.

A significant contribution by Ruan et al. can be found in this collection. In this paper, the authors aimed at finding useful *Fusarium* head blight (FHB, *Fusarium graminearum* Schwabe [teleomorph: *Gibberella zeae* (Schwein.) Petch]) resistance in durum wheat. The researchers used 186 diverse durum wheat lines, comprised of elite Canadian cultivars, breeding lines and experimental durum wheat lines with FHB resistance. The authors found 31 QTL across all durum wheat chromosomes and one stable QTL of large effect. Also, three haplotypes of the QTL *Fhb1* were identified. This large number of signals provides a treasure trove of resources for improving FHB resistance, including durable FHB resistance.

Using a core collection from the United States Department of Agriculture, Shi et al. detected signals and implementing genomic selection toward soybean cyst nematode (SCN, *Heterodera glycines*) resistance in common bean (*Phaseolus vulgaris* L.). The authors used 315 accessions from the core collection and screened for SCN. The core set was genotyped with Infinium BeadChips, consisting of 4k SNPs. A total of 15 accessions were found as resistant and 11 MTA were found. Additionally, the authors applied genome-wide prediction models and reported moderate accuracies for resistance to SCN, indicating the feasibility of using this framework when improving SCN resistance.

GENOMIC SELECTION

A significant study carried out by Merrick et al. was published. The authors conducted research to optimize GS models related to both major and minor genes for resistance to stripe rust (*Puccinia striiformis* Westend. f. sp. *tritici* Erikss.) of wheat. The authors used two types of training populations composed of 2,630 breeding lines and 475 diversity panel lines, both groups were phenotyped for 4 years. Model comparisons were also conducted using major gene markers and genome-wide markers as fixed effects. Using 50 replications and a five-fold cross-validation, the models were then compared to marker-assisted selection (MAS). The authors found that GS had higher accuracies than MAS (0.72) for disease severity. In contrast, GS and major gene models did not outperform the base GS model. Different combinations of traits, population types and years resulted in increases in accuracy as well via the inclusion of major markers in the validation sets. As well, adding fixed effects under low prediction scenarios increased GS accuracy when using significant GWAS markers. This study is a significant step in the implementation of breeding efforts for improvement of QDR.

One noteworthy contribution was provided by Larkin et al. The authors used GS for forward prediction and compared naïve GS models (no covariates) and multi-trait GS (MTGS) models by predicting $F_{4:7}$ lines for FHB resistance traits, deoxynivalenol (DON) accumulation and other traits in soft red winter wheat. They compared predictions with phenotypic performance over 2 years of selection based on selection accuracy and response to selection. The models correctly selected up to 70.1% of elite individuals, compared to 33% with phenotypic selection. The authors also measured realized response to selection for the various traits and found GS models were at least comparable to phenotypic selection for FHB. This study provides a way forward for the implementation of GS toward breeding for QDR in wheat.

Huster et al. conducted GWAS on a diversity panel of 149 snap bean pure lines and evaluated them for *Fusarium* root rot and multiple root morphological traits. The authors found five SNPs for disease severity and two for biomass, with multiple biochemical functions indicated. Genomic estimated breeding values (GEBV) were estimated across all bean lines and their correlations estimated for the development of GS models. Although low accuracies were found based on correlations, some overlap was found among lines with high GEBV and root rot resistance.

A notable submission was provided by Mphahlele et al. in the search for quantitative resistance to *Leptocybe invasa* gall wasp and fungal stem diseases such as *Botryosphaeria dothidea* and *Teratosphaeria zuluensis*. The authors deployed the *Eucalyptus grandis* EUChip60K SNP chip, a subset of 964 trees from 93 half-sib families genotyped with 14,347 SNPs. Single-step genomic best linear unbiased predictors (ssGBLUP) were used to predict parameters in the trial. The authors found a high positive genetic correlation with gall wasp tolerance moderate expected gains for traits such as diameter growth and gall wasp. This study may set future strategies for the improvement of *Eucalyptus* using GS.

MARKER DEVELOPMENT

Peach gummosis has been reportedly caused by *Botryosphaeria Fusicoccomaesculi*, *Botryosphaeria rhodina* (anamorph *Lasiodiplodia theobromae*), and *Botryosphaeria obtuse* (anamorph *Diplodiaseriata*). In their study, Li, X., et al. used a previously identified QTL from a biparental population and integrated it with a GWAS and comparative transcriptome sequencing across 195 accessions and 145 k SNPs. The authors found five SNPs linked with gummosis disease resistance and located six candidate genes in the vicinity of significant SNPs. The authors also identified two highly resistant accessions as potential sources for breeding. Cucumber vein yellowing virus (CVYV) does not exhibit single gene resistance in cucumber and is transmitted by the whitefly (*Bemisia tabaci*). Due to the lack of tightly linked molecular markers, breeding for CVYV is challenging.

A study conducted by Kahveci et al. revealed that, via the use of genomics and bulk segregant analysis, KASP markers were developed for resistance to CVYV in an F_2 mapping population and commercial lines. The authors also conducted variant analysis to generate SNP-based markers, and this resulted in a 101 kb-fine mapped region with eight putative candidate genes. Thus, the study provided crucial information and tools necessary to breed for CVYV resistance in the future.

PATHOGEN-ENVIRONMENT-GENOTYPE

In a study to understand the influence of elevated temperatures on resistance against phoma stem canker (*Leptosphaeria maculans*) in oilseed rape, Noel et al. investigated effects of temperature on individuals with and without race-specific resistance (*R*) genes and quantitative resistance. The experiments involved field sites and inoculation assays under controlled conditions and found that high maximum temperatures in June increased canker severity while this impact was reduced in genotypes with quantitative resistance but no *R* genes. This study suggested that the impact of high temperature is significantly reduced when quantitative resistance is present. The authors point out that there is genetic variation available to improve disease resistance under this condition. However, sustained high temperatures reduce the efficacy of QDR—a major concern in the face of global warming/climate change.

In Ozimati et al., the authors evaluated empirical and root necrosis data to determine the effectiveness of screening for Cassava brown streak disease (CBSD) in two breeding populations differing in selection cycles. The study aimed at comparing the assessments in these screening methods when the assessment was conducted by plant breeders vs. pathologists. The study found that broad-sense and marker-based heritability estimates differed widely from assessments within the two groups, with breeders resulting in a slightly higher upper limit.

BREEDING AND PRE-BREEDING

Emebiri et al. reported efforts on pre-emptive breeding for Karnal bunt in wheat. Karnal bunt caused by the fungus, *Tilletia*

indica Mitra [syn. *Neovossia indica* (Mitra) Mundkur], is a major threat to food security, due to its use as a non-tariff trade barrier by several wheat-importing countries. The cultivation of resistant varieties remains the most cost-effective approach to manage the disease, but in countries that are free of the disease, genetic improvement is difficult due to quarantine restrictions. Using GWAS, the authors identified six DArTseq markers linked with resistance to Karnal bunt, each marker explained up to 29.5% of phenotypic variation. Using GS, the authors reported accuracies of up to 0.56, depending on whether the GS model included known QTL or used genome-wide markers. The authors conducted further research to identify elite parents with Karnal bunt resistance, leading to the identification of one ideal genotype with suitable agronomic traits.

The study in Awata et al. aimed at using backcross populations at the BC₃F₂ generation to identify progeny resistant to maize lethal necrosis (MLN) developed using marker-assisted backcrossing. The research group used SNP based markers linked to six QTL for resistance through screening 2,400 BC₃F₂ lines using a KASP platform. The authors found 56 BC₃F₂ lines that had major resistance for MLN and confirmation experiments in the field resulted in 19 lines with high levels of MLN resistance. These validated KASP markers linked to the two major QTL which will serve to speed up breeding.

Ballén-Taborda et al. used wild relatives of peanut to conduct marker-assisted backcrossing of two loci controlling resistance to peanut root-knot nematode (PRKN, *Meloidogyne arenaria*) from *Arachis stenosperma*. The study performed four cycles of backcrossing while utilizing SNP analysis for foreground selection. A population of 271 BC₃F₁ lines was genotyped to determine introgression level across the peanut genome. The results indicate that PRKN resistance was validated in BC₃F₃ lines with seed size characteristics maintained. The study concluded that it is the introgression of both loci validated from *A. stenosperma* that confer the resistance. The work will represent a significant step in breeding for PRKN resistance into elite peanut cultivars.

REVIEWS

Soares et al. reviewed current progress in the improvement of disease resistance to *Mycosphaerella fijiensis* Morelet [anamorph: *Pseudocercospora fijiensis* (Morelet) Deighton] in bananas (*Musa* spp.). With the use of pre-established exclusion and inclusion criteria, the paper is a systematic review of papers collected using six scientific journal databases analyzing 3,070 published studies, identifying 24 relevant to the *Musa-M. fijiensis* pathosystem. Relevant articles found revealed that variable response to sigatoka exists among resistant and susceptible cultivars. In the case of *M. acuminata* wild diploids, resistance genes exist, and these have served as parental for new generations of improved diploids and introgression into elite cultivars. One of the highlights of the review indicates that the sequencing of the resistance genes in the *M. acuminata* genome still require functional validation across multiple omics data layers. Previously reported resistance genes have been involved in primary disease resistance pathways, such

as jasmonic acid and ethylene signaling. Gene-based markers have been reported in *Musa* and are applicable for MAS. This review is a comprehensive panorama of the immune response found in the *Musa-P. fijiensis* pathosystem and summarizes some of the avenues available for breeders to undertake efforts to develop resistant cultivars.

Manze et al. provided an overview of genetic gain yield and virus resistance in Cassava over the last eight decades in Uganda. This study used 32 Cassava varieties released between 1940 and 2019 and conducted side-by-side multilocation trials in Uganda. Although disease resistance increased at an average up to 2.3% per year, fresh root yield and harvest index genetic gains have been non-significant. The authors reported some of the progress made in Cassava breeding, as well as some of the challenges that have yet to be solved, highlighting that breeding has mainly focused on protecting cassava against diseases while agronomic performance has not received sufficient attention.

FINAL REMARKS

Papers published in this Research Topic highlight progress made in classical and modern breeding for QDR, a trait that used to be evaluated solely based on visual symptom rating. It is known that QDR occurs probably in all pathosystems and is caused by the presence of multiple loci distributed across the whole genome of plants. The effect of each loci varies ranging from small to large and can be affected by environmental factors and plant growth stages as reported in this Topic. Additionally, quantitative disease loci (QDL) appears to be independent from the presence of qualitative resistance; thus, it is possible that breeding for QDR may not necessarily include genes for qualitative resistance. With the use of QTL and GWAS approaches, QDR can be analyzed by bi- or multi-parental QTL mapping and/or GWAS on diversity panels to identify QRLs in host plant genome that are closely associated with QDR. Dissecting QRLs could allow to isolate genes responsible for QDR. From breeding perspective, QRLs can be used for MAS to effectively develop new cultivars with durable resistance to pathogens of interest. The implementation of new tools (e.g., genomic selection) that enable accurate selection of large numbers of marker haplotypes simultaneously is a promising avenue for the accumulation of favorable alleles contributing to QDR. However, detecting and managing Genotype x Environment interaction in breeding for QDR continues to be a challenge.

AUTHOR CONTRIBUTIONS

VH-V: wrote the editorial. VH-V, JC, AM, and HR: served as lead topic editor for the collection. JC, AM, and HR: assisted with writing and proofreading the editorial. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The topic editors wish to thank all the contributing authors to this collection.

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 Hoyos-Villegas, Chen, Mastrangelo and Raman. 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 of the Genetic Architecture for Fusarium Head Blight Resistance in Durum Wheat: The Complex Association of Resistance, Flowering Time, and Height Genes

Yuefeng Ruan¹, Wentao Zhang^{2*}, Ron E. Knox¹, Samia Berraies¹, Heather L. Campbell¹, Raja Ragupathy³, Kerry Boyle², Brittany Polley², Maria Antonia Henriquez⁴, Andrew Burt⁵, Santosh Kumar⁶, Richard D. Cuthbert¹, Pierre R. Fobert⁷, Hermann Buerstmayr⁸ and Ron M. DePauw^{9,10}

¹ Swift Current Research and Development Centre, Agriculture and Agri-Food Canada, Swift Current, SK, Canada, ² Aquatic and Crop Resource Development Research Centre, National Research Council of Canada, Saskatoon, SK, Canada, ³ Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, AB, Canada, ⁴ Morden Research and Development Centre, Agriculture and Agri-Food Canada, Morden, MB, Canada, ⁵ Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, ON, Canada, ⁶ Brandon Research and Development Centre, Agriculture and Agri-Food Canada, Brandon, MB, Canada, ⁷ Aquatic and Crop Resource Development Research Centre, National Research Council of Canada, Ottawa, ON, Canada, ⁸ University of Natural Resources and Life Sciences, Vienna, Austria, ⁹ Advancing Wheat Technology, Swift Current, SK, Canada, ¹⁰ Retired from Swift Current Research and Development Centre, Agriculture and Agri-Food Canada, Swift Current, SK, Canada

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Ken Chalmers,
University of Adelaide, Australia
Angelica Giancaspro,
University of Bari Aldo Moro, Italy

*Correspondence:

Wentao Zhang
Wentao.Zhang@nrc-cnrc.gc.ca

Specialty section:

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

Received: 06 August 2020

Accepted: 30 November 2020

Published: 23 December 2020

Citation:

Ruan Y, Zhang W, Knox RE, Berraies S, Campbell HL, Ragupathy R, Boyle K, Polley B, Henriquez MA, Burt A, Kumar S, Cuthbert RD, Fobert PR, Buerstmayr H and DePauw RM (2020) Characterization of the Genetic Architecture for Fusarium Head Blight Resistance in Durum Wheat: The Complex Association of Resistance, Flowering Time, and Height Genes. *Front. Plant Sci.* 11:592064. doi: 10.3389/fpls.2020.592064

Durum wheat is an economically important crop for Canadian farmers. Fusarium head blight (FHB) is one of the most destructive diseases that threatens durum production in Canada. FHB reduces yield and end-use quality and most commonly contaminates the grain with the fungal mycotoxin deoxynivalenol, also known as DON. Serious outbreaks of FHB can occur in durum wheat in Canada, and combining genetic resistance with fungicide application is a cost effective approach to control this disease. However, there is limited variation for genetic resistance to FHB in elite Canadian durum cultivars. To explore and identify useful genetic FHB resistance variation for the improvement of Canadian durum wheat, we assembled an association mapping (AM) panel of diverse durum germplasms and performed genome wide association analysis (GWAS). Thirty-one quantitative trait loci (QTL) across all 14 chromosomes were significantly associated with FHB resistance. On 3BS, a stable QTL with a larger effect for resistance was located close to the centromere of 3BS. Three haplotypes of *Fhb1* QTL were identified, with an emmer wheat haplotype contributing to disease susceptibility. The large number of QTL identified here can provide a rich resource to improve FHB resistance in commercially grown durum wheat. Among the 31 QTL most were associated with plant height and/or flower time. QTL 1A.1, 1A.2, 3B.2, 5A.1, 6A.1, 7A.3 were associated with FHB resistance and not associated or only weakly associated with flowering time nor plant height. These QTL have features that would make them good targets for FHB resistance breeding.

Keywords: resistance, QTL, GWAS, Fusarium head blight, durum

INTRODUCTION

Fusarium head blight (FHB), also known as scab and mainly caused by *Fusarium graminearum* Schwabe [teleomorph: *Gibberella zeae* (Schwein.) Petch] (Bai and Shaner, 1994; McMullen et al., 1997), is a devastating fungal disease of small-grain cereals including durum and common wheat and barley, resulting in severe yield and quality losses (Gilbert and Tekauz, 2000; McMullen et al., 2012). Moreover, as food for humans and feed for animals, FHB infected grain also creates health risks due to contamination with mycotoxins. This is a particular concern for durum wheat, as its main purpose is for human consumption (Bai and Shaner, 2004; Zhao et al., 2018; Haile et al., 2019; He et al., 2019). Canada is the largest producer and exporter of durum wheat supplying more than a half of the world's total exported durum (International Grains Council, 2020). Since the early 1990s, FHB has become the major disease threatening durum production in Canada and has caused major economic losses for producers (Gilbert and Tekauz, 2000). In 2016, a severe FHB epidemic caused 65% of the common wheat and 36% of the durum wheat to be downgraded in Saskatchewan, Canada, with an estimated economic loss of \$1 billion (Canadian Grain Commission, 2017). It is therefore a priority to develop durum wheat with desirable FHB resistance to protect it from losses.

Currently, the combination of agronomic and chemical control along with genetic resistance is the most effective means to manage FHB (Gilbert and Haber, 2013; Prat et al., 2014). Genetic resistance is preferred due to its lower cost, higher efficacy, and environmental benefit (Prat et al., 2014). Genetic resistance to FHB in wheat is quantitative in expression due to control by multiple minor genes. FHB resistance is also significantly affected by environment (Bai and Shaner, 2004; Buerstmayr et al., 2009, 2019), thus having lower to moderate heritability (Van Sanford et al., 2001). Therefore, when visual assessment of FHB is performed in the field, lines must be tested in multiple independent environments with intensive phenotyping to reliably identify QTL for resistance.

Developmental traits including flower time, plant height, spike morphology, and anther extrusion/or retention are often reported for their relationship with FHB resistance (Mesterhazy, 1995; Gervais et al., 2003; Srinivasachary et al., 2009; Skinnies et al., 2010; Lv et al., 2014; Buerstmayr and Buerstmayr, 2016). Plant height and disease resistance mostly show a significantly negative correlation (Mesterhazy, 1995; Srinivasachary et al., 2009; Buerstmayr and Buerstmayr, 2016). Pleiotropic effects, tightly linked genes and disease escape have all been hypothesized as feasible mechanisms for resistance related to these developmental traits.

Fusarium head blight resistance can be categorized into three main types or components: (1) type I – resistance to initial infection measured by the incidence of disease in the presence of natural or augmented artificial inoculum (e.g., spray inoculation); (2) type II – resistance to fungal spread measured by the severity of disease; and (3) type III – resistance to the accumulation of the toxin deoxynivalenol (DON) in infected spikes (Miller et al., 1985; Mesterhazy, 1995; Bai and Shaner, 2004). Till now, more than 556 QTL contributing to FHB resistance have

been identified on all 21 chromosomes of hexaploid wheat (Buerstmayr et al., 2009; Liu et al., 2009; Löffler et al., 2009; Venske et al., 2019). These QTL can be refined largely into 56 clusters by meta-QTL analysis (Venske et al., 2019). In spite of the relatively large number of identified QTL for FHB, only three QTL, *Fhb1* on chromosome arm 3BS (Anderson et al., 2001; Liu et al., 2006), *Qfhs.ifa-5A* on 5AS (*Fhb5*) (Buerstmayr et al., 2002; Somers et al., 2003; Steiner et al., 2019a) and *Fhb2* on 6BS (Anderson et al., 2001; Cuthbert et al., 2007) have been validated. All of these resistance loci originate from the Chinese cultivar Sumai 3, which displays among the highest levels of FHB resistance observed (Buerstmayr et al., 2009). *Fhb1* is the best validated, and most frequently studied and deployed resistance QTL (Buerstmayr et al., 2019). It is currently the only resistance QTL confirmed to be present in several new FHB North American and European varieties with strong resistance (Hao et al., 2019). *Fhb1* is reported primarily as conferring strong Type II resistance, and accounting for 20–60% of phenotypic variation in breeding populations (Miedaner and Korzun, 2012). *Fhb1* was recently claimed to be cloned by three research groups as two different candidate genes (Rawat et al., 2016; Li et al., 2019; Su et al., 2019) with conflicting interpretations, leaving room for independent validation.

Compared to the large amount of genetic variation for FHB resistance reported in common wheat, durum wheat has limited sources of resistance (Oliver et al., 2008; Prat et al., 2014; Steiner et al., 2019b). Tetraploid sources of FHB resistance that have been identified include the Canadian durum cultivar Strongfield (Somers et al., 2006), experimental line DT696 (Sari et al., 2018), *T. carthlicum* (Somers et al., 2006; Oliver et al., 2008; Sari et al., 2018), *T. dicoccoides* (Ruan et al., 2012), *T. dicoccum* (Buerstmayr et al., 2012; Zhang et al., 2014), and Tunisian durum landraces (Ghavami et al., 2011; Huhn et al., 2012). Among these findings, the most stable and consistent QTL were identified on chromosomes 2B, 3A, 3B, and 5A (Prat et al., 2014; Haile et al., 2019).

As hexaploid wheat has significantly more sources of FHB resistance, introgression of resistance from hexaploid into durum wheat is one possible way to expand the durum resistance gene pool. Previous attempts to introgress FHB resistance from Sumai 3 into durum were largely unsuccessful (Prat et al., 2017). However, several recent successes have been reported with *Fhb1* from Sumai 3 (Giancaspro et al., 2016; Prat et al., 2017) as well as a non-Sumai 3-related FHB resistance sources (Chu et al., 2011; Zhao et al., 2018). Despite these partial successes, no commercial durum cultivars with QTL from these non-adapted sources have been released due to the lengthy breeding process, linkage drag or suppression of resistance in durum backgrounds. Because of these challenges, utilizing the FHB resistance already present in durum cultivars is gaining favor as a promising approach to bring durum wheat cultivars with improved resistance to market more quickly. Durum cultivars with an improved level of FHB resistance have been developed and released by the North Dakota durum breeding program using this strategy (Zhang et al., 2014). With the same approach, recent durum cultivars, including Brigade (Clarke et al., 2009) and Transcend (Singh et al., 2012) with a better level of FHB resistance have also been

successfully developed and released by Canadian durum breeding programs selecting for reduced symptoms in FHB nurseries. Regardless of this initial success, there is still a need to know and identify additional native sources of resistance as well as more exotic sources. Understanding the association of FHB resistance with developmental traits, flowering time and plant height is also important for recommending which resistance loci may be most relevant to a particular breeding program.

Genome wide association studies (GWAS) are a promising way to detect FHB resistance QTL present in diverse genetic sources. Only a few GWAS have been conducted on FHB resistance, including winter wheat (Wang et al., 2017), elite Chinese wheat (Zhu et al., 2020), durum breeding panels (Steiner et al., 2019b) and type II FHB resistance durum diversity panels (Ghavami et al., 2011). In this study, we aimed to use GWAS to explore FHB resistance of domestic durum cultivars and breeding material as well as exotic sources of resistance, including Sumai 3 and emmer wheat introgression lines. With GWAS in multiple environments, we aimed to: 1) explore and characterize FHB resistance QTL in durum wheat from the domestic as well as exotic sources, and 2) identify resistance QTL that colocalize with flowering time and plant height.

MATERIALS AND METHODS

Plant Materials

In total, 186 diverse durum wheat (*Triticum turgidum* L. ssp. *Durum* (Desf.) Husn.) lines were selected to constitute a durum association mapping (AM) panel targeted to improve FHB resistance in durum wheat. This panel was primarily composed of durum from Canada, including elite Canadian cultivars, advanced breeding lines, recently developed germplasm from Canadian breeding programs and from research projects (Supplementary Table 1). Experimental durum lines representing exotic FHB resistance and germplasm from global collections made up the remainder of the AM panel (Supplementary Table 1).

Phenotyping

Lines of the durum AM panel were evaluated for FHB infection in Morden and Brandon, MB, Canada in 2015 to 2017 with artificial inoculation and Indian Head, SK, Canada in 2015 and 2016 with natural infection. At both Morden and Brandon, FHB nurseries, corn spawn inoculum of *Fusarium graminearum* was used. Corn spawn consisted of grains that were inoculated with a mixture of two *F. graminearum* isolates, a 3-acetyl-deoxynivalenol (3ADON, M9-07-01) and a 15-acetyl-deoxynivalenol (15ADON, M1-07-02) isolate, after which colonized kernels were air dried. In Morden, approximately 2–3 weeks prior to heading, the corn spawn inoculum was spread at 8 g per single meter row with two applications at weekly intervals. Plots were irrigated three times per week using Cadman Irrigation travelers with Briggs booms. At Brandon, the corn spawn inoculum was applied between the rows at a rate of 40 g/m 6 weeks after planting, with a second application performed at the same rate 2 weeks after the first. Plots were irrigated three times

per week with a mist irrigation system to create favorable conditions for *F. graminearum* infection. In Indian Head, FHB was achieved solely by natural disease infection. FHB incidence (INC, percentage of spikes showing symptoms) and severity (SEV, average percentage of spike with visual symptoms of infection) were estimated with visual assessment. FHB index (IND) was calculated with the formula: $(INC \times SEV)/100$. Plant height (HT) and days to anthesis (DTA) were also recorded for Morden plots.

Genotyping

Genomic DNA of the durum AM panel was extracted from freeze-dried fresh leaf tissue of seedlings with a CTAB based protocol carried out on an automated AutoGen DNA isolation system (AutoGen, Holliston, MA). DNA was quantified with a Quant-iTTM PicoGreen[®] dsDNA Assay Kit (Thermo Fisher Scientific Inc., Bartlesville, OK, United States) and diluted to 50 ng/μL for SNP array genotyping. Genotyping of DNA was performed with the Illumina iSelect 90K SNP array (Wang et al., 2014) according to the manufacturer protocol (Illumina). SNP arrays were scanned with an Illumina HiScan. Raw intensity files from the HiScan were imported into GenomeStudio Version 2013 (polyploid clustering module v1.0.0, Illumina). SNP calling was performed with the method described by Wang et al. (2014) with 3 cluster steps of the cluster algorithm DBSCAN then OPTICS. All SNPs were subsequently visually checked, and incorrectly clustered SNPs or SNPs with more than 4 clusters were manually removed. Finally, SNPs with minor allele frequency (MAF) below 0.05, and missing genotypes higher than 15% were filtered out. This resulted in a total of 6900 high quality polymorphic SNPs of which 5933 markers were anchored to the wheat consensus map (Wang et al., 2014) for the downstream genome wide association analysis.

Statistical Analysis

Statistical analysis was performed with R 3.4.2 (R Core Team, 2017) with the lme4 package (Bates et al., 2015). Phenotypic traits from each disease nursery site across multiple years were fitted with the linear mixed model (Bates et al., 2015). The model is implemented as: $P_{iy} = \mu + G_i + E_y + (G_i \times E_y) + E_{iy}$, where, P_{iy} are the values of the tested phenotypic trait, μ is the population mean, G_i is the effect of genotypes, E_y is the effect of environments (here, by Year), E_{iy} is the residual, where i is the genotype, y is the year. The restricted maximum likelihood (REML) method within lme4 (Bates et al., 2015) was used to estimate the variance components of each trait. The broad sense heritability (H^2) was estimated with the equation

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma_{G \times E}^2}{y} + \frac{\sigma_e^2}{p}}$$

site, where: σ_G^2 is the genotypic variance, $\sigma_{G \times E}^2$ is the variance of interaction between genotype and year, σ_e^2 is the error variance, y is the number years, and p is the total number of replications in all tested years. The least squares means were used for trait correlation and association mapping analysis. The correlation coefficients of disease response, plant height and days to anthesis across multiple years and multiple sites were calculated with

the Pearson correlation test and visualized with the R package “corrplot” (Wei and Simko, 2017).

Linkage Disequilibrium, Population Structure, and Kinship Analysis

Linkage disequilibrium (LD) was estimated by correlation coefficient analysis and used the squared correlation coefficients (r^2) for all 5933 anchored SNP markers implemented in Tassel v.5.5.0 (Bradbury et al., 2007). The r^2 values of unlinked genetic markers (defined as genetic distance > 30 cM) were square-root transformed into a normal distribution. The baseline (or critical) r^2 value, a value that suggested LD was likely caused by genetic linkage, was determined by taking the 95% percentile of this distribution (Brescaghello and Sorrells, 2006). The scatter plot of r^2 versus genetic distance (cM) was fitted using a non-linear model described by Remington et al. (2001) that was implemented in software PopLDdecay (Zhang et al., 2019).

Population structure of the durum AM panel was determined with STRUCTURE v2.3.4 (Pritchard et al., 2000) with a pruned SNP marker dataset that was generated with the LD (linkage disequilibrium)-based pruning approach implemented in the software PLINK (Purcell et al., 2007). A total of 2306 pruned markers with LD (r^2) \leq 0.2 were used for population structure analysis. STRUCTURE analysis was performed with a 50000 burn-in length and 100000 Markov chain Monte Carlo (MCMC) iterations from $K = 2$ to $K = 12$ (K , specialized clusters of the AM panel). Fifteen independent STRUCTURE runs were conducted for each specialized K . The optimal cluster (K) was determined by the ΔK method (Evanno et al., 2005), implemented in the software Structure Harvester (Earl and vonHoldt, 2012). Independent runs of the optimal K were summarized using CLUMMP v1.1.2 software (Jakobsson and Rosenberg, 2007). The CLUMMP generated Q matrix was used to graph the population structure using Structure Plot software (Ramasamy et al., 2014) and perform downstream GWAS analysis. A phylogenetic tree was built with the neighbor-joining (NJ) method in MEGA6 (Tamura et al., 2013) and visualized with Figtree v1.4.4¹. Principal component analysis (PCA) was performed with R package Genome Association and Prediction Integrated Tools (GAPIT) (Lipka et al., 2012; Tang et al., 2016).

Genome-Wide Association Study

Association mapping was performed on the durum association panel using the phenotypic data collected from the multiple nurseries in multiple years, including HT, DTA, INC, SEV and IND. Association mapping was performed using 5933 mapped SNPs that had a MAF > 0.05 using both Tassel v5.5.50 (Bradbury et al., 2007) and GAPIT (Lipka et al., 2012; Tang et al., 2016). Different association models were tested in both software packages, and QQ-plots generated from all the models were compared to select the model that best controls false positives and negatives. All the data presented here were generated in TASSEL using a mixed linear model (MLM) incorporating the STRUCTURE (Q) matrix as a fixed factor and the kinship (K) matrix as a random factor (Q + K MLM). To be considered a QTL in this dataset, we selected SNPs that were significant

($p < 0.05$, marker-wise) in at least four of the tested environments for FHB resistance or two for plant height and flower time, and with at least one environment with a highly significant response ($p < 0.001$). Significant SNPs on the same linkage group were grouped into a QTL region if markers were linked with LD > 0.2.

RESULTS

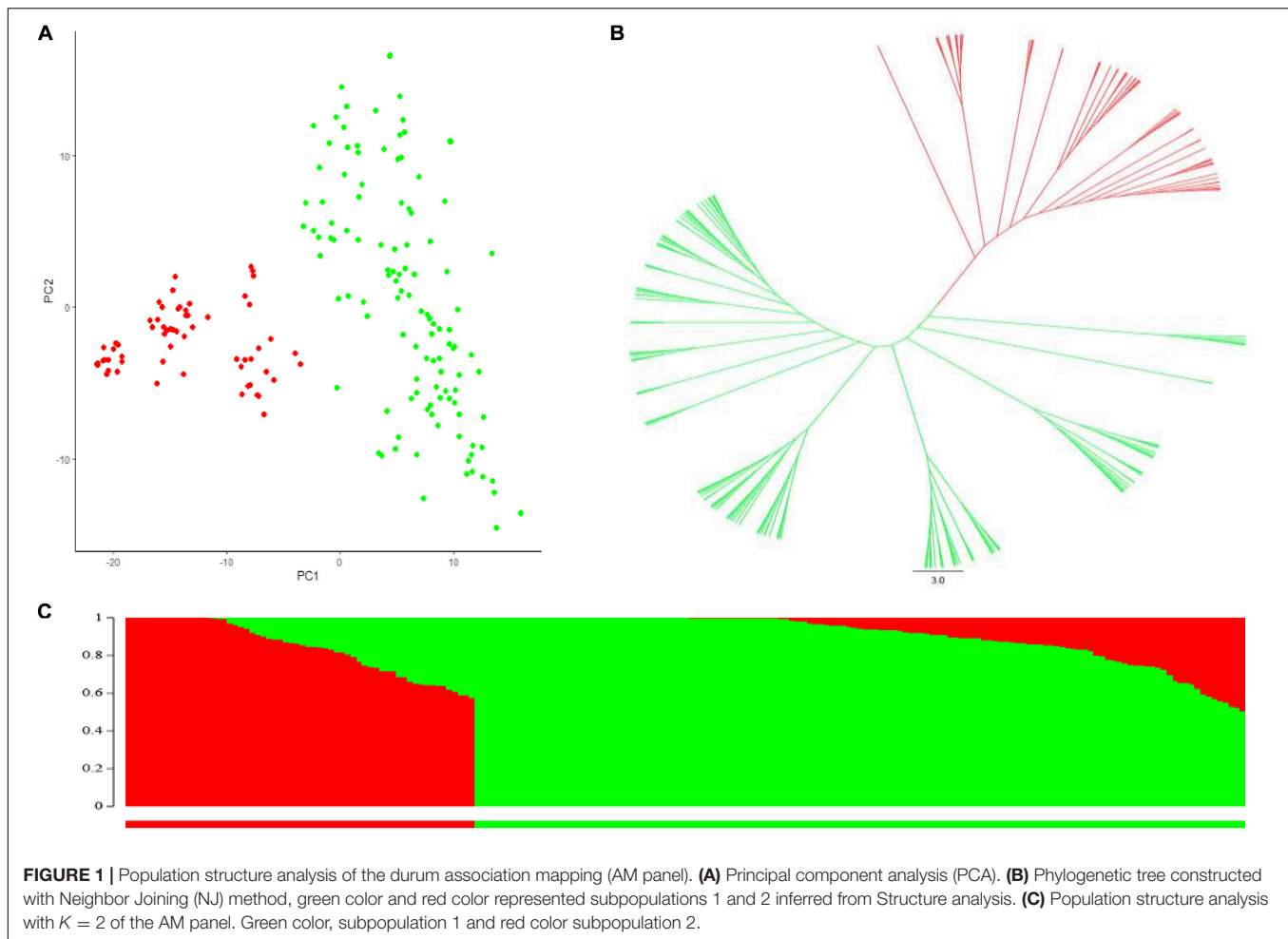
Population Structure and Linkage Disequilibrium (LD) Analysis

STRUCTURE analysis, principal component analysis (PCA) and NJ-phylogenetic tree analysis were all used to determine clustering of lines within the durum AM panel, and two subpopulations were consistently indicated, as shown by different colors in **Figure 1**. Subpopulation 1 (shown as green in **Figure 1** and **Supplementary Table 1**) contained 124 lines, and consisted of a large proportion of Canadian cultivars and inbred lines including the older cultivar Kyle, more recent cultivars Strongfield and currently most popular cultivars as Brigade, Transcend and CDC Credence. Subpopulation 2 included 62 lines (shown in red in **Figure 1** and **Supplementary Table 1**), consisting of the founder landrace Pelissier and the majority of lines from Austria. All of the inbreeding lines derived from introgression of FHB resistance genes from Sumai 3 into European durum wheat cultivars were contained in subpopulation 2, as were the majority of *T. dicoccoides* introgression lines. The baseline critical threshold r^2 value of LD was identified as 0.2, corresponding to a genetic distance around 3.0 cM from the whole genome analysis (**Supplementary Figure 1**).

Phenotypic Analysis

Mean values (across years) of FHB INC, SEV, IND, DTA and HT of lines from the durum AM panel at Brandon, Morden, and Indian Head, were summarized in **Supplementary Table 1**. Across environments, FHB INC tended to be higher than SEV (**Figure 2**) which is reflected in the overall means (**Table 1**). The lowest INC was observed at Indian Head in 2016, the location with the lowest severities in both 2015 and 2016. Moderate SEV were observed at Brandon in 2016 and 2017. Generally, a large differential in FHB INC and SEV was observed as indicated by the range for each environment in **Table 1**, except Indian Head where the maximum severity of disease was less than 100%. Plant height showed a larger range with the average shortest 55 cm and the highest 148 cm while DTA was observed in a range of 13 days in 2017 and 20 days in 2015 (**Table 1** and **Supplementary Figure 2**). For both INC and SEV, moderate to high broad sense heritability was observed with the two sites under artificial inoculation showing lower heritability than the natural infection site (**Table 1**). HT showed the highest heritability, while DTA had the lowest heritability (**Table 1**). For FHB INC and SEV, moderate to high correlations were observed in all tested environments (years and sites). Generally, both HT and DTA had very significant negative correlations with INC and SEV (**Figure 3**). Analysis of variance (ANOVA) revealed

¹ <http://tree.bio.ed.ac.uk/software/figtree/>



that genotypic effects were significant for all phenotypic traits ($P < 0.001$, **Supplementary Table 2**).

GWAS Analysis of FHB Resistance, HT and DTA

With GWAS analysis, 31 genomic regions were significantly associated with FHB resistance traits (**Figures 4, 5**). The quantile-quantile (QQ) plots (**Supplementary Figure 3**) showed that, for the majority of traits, an appropriate model was fitted for the GWAS test. The GWAS results were summarized in **Table 2** and **Supplementary Table 3**. SNPs located within the same region were grouped into QTL, and **Table 2** shows the QTL names and physical location of the associated SNPs based on their location on the IWGSC Chinese Spring (CS) reference 1.0 (CS Ref 1.0; International Wheat Genome Sequencing Consortium [IWGSC], 2018). For each significant QTL, the lowest $-\log_{10}(p\text{-value})$ is shown for each environment and trait tested whenever the p -value is less than $p = 0.05$. As shown in **Table 2**, there was significant variation in detection of QTL across all of the environments, and more detection of INC than SEV across the environments. The majority of the FHB resistance QTL colocalized with DTA and/or HT.

A major QTL, *1B.1*, was found between 544 and 580 Mb on 1B (**Figure 5** and **Table 2**). It was significant for INC, SEV and IND, and explained as much as 20% of the phenotypic variation (**Table 2** and **Supplementary Table 3**). The QTL *1A.3* was located in the syntenic region of *1B.1*, between 503 and 580 Mb (**Figure 5** and **Table 2**), and it was also significant for IND and INC though present in fewer environments and with lower significance than *1B.1* (**Table 2**). *1B.1* colocalized with significant HT and DTA QTL, while *1A.3* was significant for DTA.

Another major QTL was at 30–31 Mb on 2AS, termed *2A.1* (**Figure 5** and **Table 2**). This QTL was significant for INC, IND and SEV, as well as being associated with HT and DTA (**Figures 4, 5**, **Tables 2**, and **Supplementary Table 3**). It was one of the more stable QTL detected, being present for INC in all environments. Another significant QTL, *2B.1*, was located between 8.6 and 22 Mb and was associated with all tested traits and explained up to 15% phenotypic variation (**Table 2** and **Supplementary Table 3**). QTL *2A.2* was also stable, and detected for INC in seven, IND in eight and SEV in five environments (**Supplementary Table 3**). It was located from 138 to 142 Mb, and was consistently associated with DTA (**Table 2** and **Supplementary Table 3**). QTL *2A.2* explained up to 10% of phenotypic variation (**Table 2**). On group 5, the QTL *5A.1*

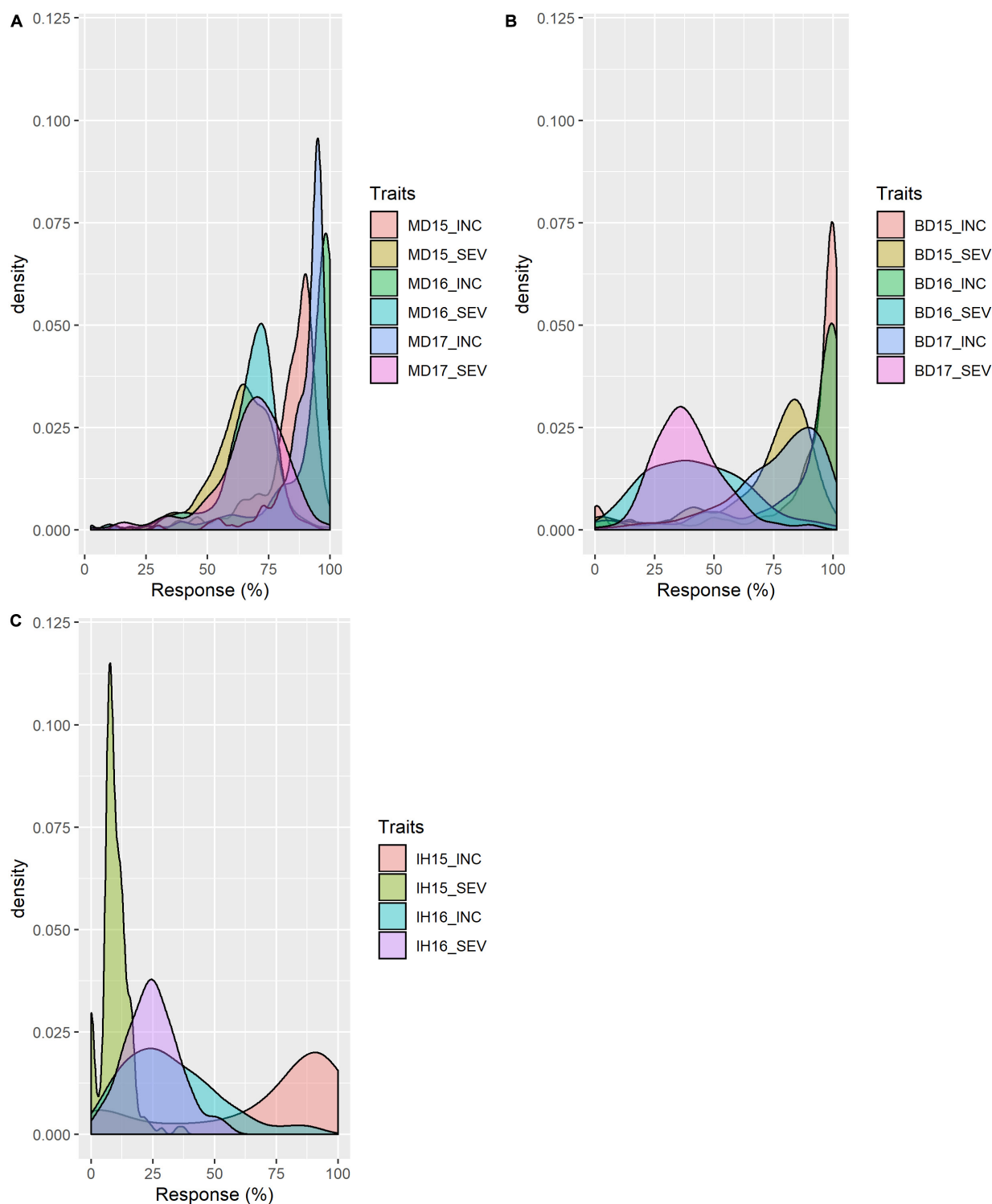


FIGURE 2 | Distribution of FHB resistance of the durum association mapping panel (AM) in field trials at **(A)** Morden, MB; **(B)** Brandon, MB; and **(C)** Indian Head, SK. INC: incidence (%), percentage of spikes showing symptoms; SEV: severity (%), percentage of spike area infected. 15, 16 and 17: years 2015, 2016, and 2017.

in the region between 585 and 591 Mb of 5A had a relatively stable effect for INC in both Brandon, MB, and Indian Head, SK (Figures 4, 5 and Table 2). It was detected at a low level for HT

in one environment. 5B.2 was located from 577 to 691 Mb on 5BL (Table 2). It explained up to 9.6% of phenotypic variation, and was most stable for INC in Brandon and Morden. This

TABLE 1 | Mean, range and heritability of the durum association mapping panel (AM) for FHB incidence, FHB severity, plant height (cm), and days to anthesis (DTA) for the individual trial in Morden, Brandon, and Indian Head across the 2015–2017 trial series, and across sites between Morden and Brandon.

| Sites | Traits | Year | Mean | Max | Min | H ² |
|--------------------|-----------------|---------|------|-----|-----|----------------|
| Morden | FHB incidence | 2015 | 83.0 | 100 | 15 | 0.82 |
| | | 2016 | 90.3 | 100 | 10 | |
| | | 2017 | 89.3 | 100 | 0 | |
| | | Overall | 87.5 | 100 | 0 | |
| | FHB severity | 2015 | 63.9 | 100 | 10 | 0.86 |
| | | 2016 | 66.8 | 100 | 10 | |
| | | 2017 | 67.3 | 100 | 0 | |
| | | Overall | 66.0 | 100 | 0 | |
| | Plant height | 2015 | 94.9 | 135 | 55 | 0.94 |
| | | 2016 | 99.6 | 148 | 62 | |
| | | 2017 | 92.4 | 136 | 55 | |
| | | overall | 96.6 | 148 | 55 | |
| | Day to anthesis | 2015 | 59.1 | 72 | 52 | 0.56 |
| | | 2016 | 62.7 | 73 | 54 | |
| | | 2017 | 67.6 | 76 | 63 | |
| | | Overall | 63.3 | 76 | 52 | |
| Brandon | FHB incidence | 2015 | 84.8 | 100 | 0 | 0.86 |
| | | 2016 | 85.0 | 100 | 0 | |
| | | 2017 | 76.2 | 100 | 0 | |
| | | Overall | 82.0 | 100 | 0 | |
| | FHB severity | 2015 | 70.7 | 100 | 0 | 0.77 |
| | | 2016 | 42.3 | 100 | 0 | |
| | | 2017 | 39.2 | 100 | 0 | |
| | | Overall | 50.7 | 100 | 0 | |
| Morden and Brandon | FHB incidence | | 84.0 | 100 | 0 | 0.72 |
| | FHB severity | | 58.4 | 100 | 0 | 0.67 |
| Indian Head | FHB incidence | 2015 | 70.6 | 100 | 0 | 0.60 |
| | | 2016 | 33.1 | 100 | 0 | |
| | | Overall | 51.6 | 100 | 0 | |
| | FHB severity | 2015 | 9.5 | 40 | 0 | 0.59 |
| | | 2016 | 22.7 | 60 | 0 | |
| | | Overall | 16.1 | 60 | 0 | |

Maximum (Max) and minimum (Min) values observed for traits, broad sense heritability coefficient (H²).

QTL was also associated with DTA, and minor effects were observed on IND and SEV, including at Indian Head (**Figure 4** and **Table 2**).

Three QTL were identified on chromosome 3B (**Figure 5**). The 3B.1 QTL was located around 3.7 Mb. It was identified in significant levels for INC, IND and SEV, and explained as much as 20.8% of phenotypic variation (**Table 2**). The QTL also affected HT with a very large effect on DTA. A stable QTL, designated 3B.3, was located on chromosome 3B at 141–233 Mb (**Table 2**). This QTL affected up to 9% of phenotypic variation, and also conferred a very stable effect on HT and smaller effect on DTA (**Table 2**). The third 3B QTL, 3B.2, was located around 9.8 Mb (**Table 2**), approximately 1 Mb from *Fhb1* in common wheat (Rawat et al., 2016; Li et al., 2019; Su et al., 2019). It had no observable effect on HT or DTA, but also

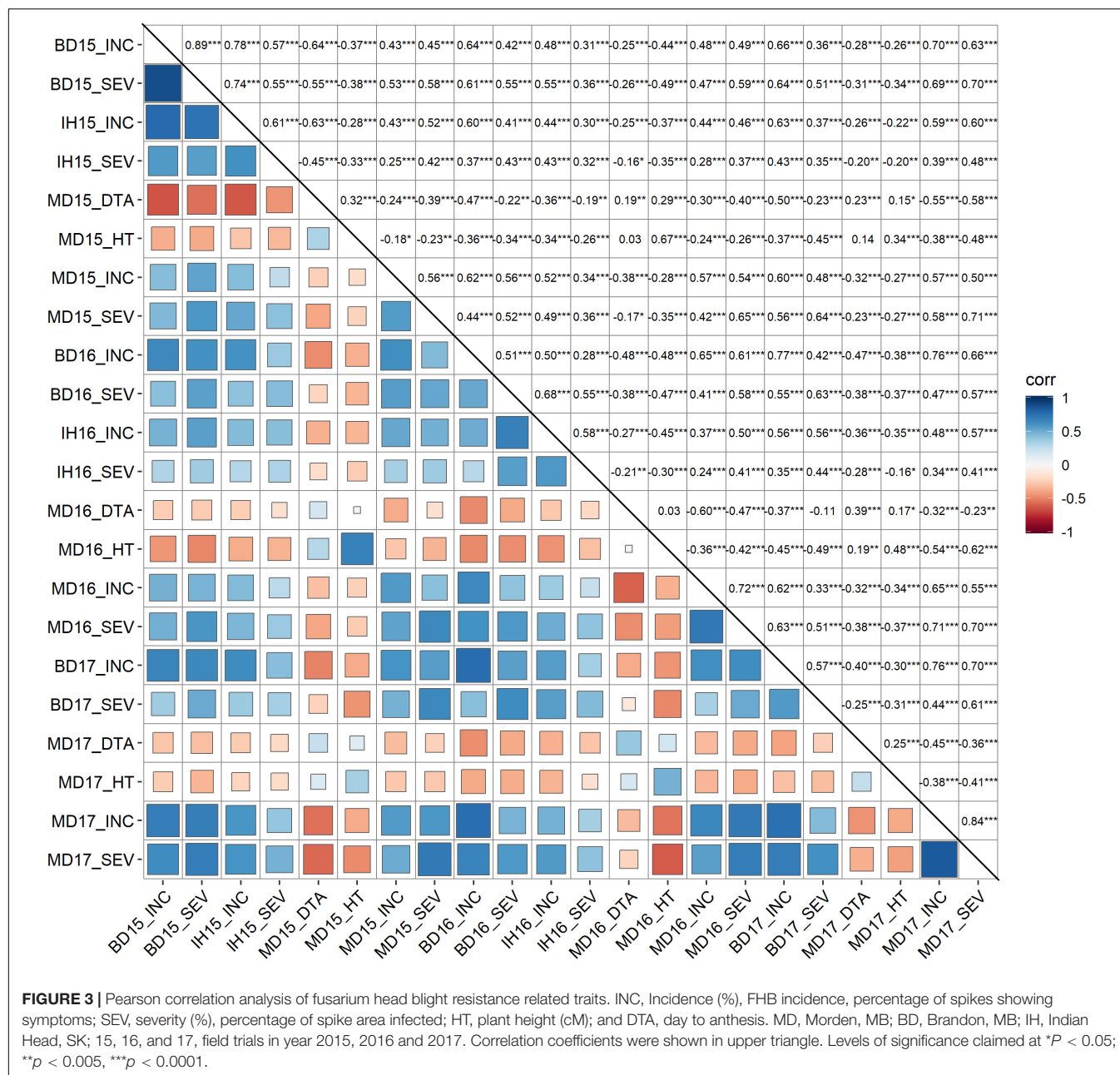
had a quite minor effect, explaining at most 7.7% of phenotypic variation (**Table 2**). Though this QTL was less stable, because of the location of 3B.2 in the region of *Fhb1* and because of the importance of this gene to FHB resistance in common wheat, we chose to further characterize the QTL in the durum AM panel. Pedigree information and genotypes of 3B.2 identified three different haplotypes for the significant marker, BS00079522_51, which were defined as tSumai3, tNative and tEmmer types (**Supplementary Table 4**). The tSumai 3 haplotype was derived from the introgression of *Fhb1* from Sumai 3 into durum wheat (**Supplementary Table 4**). All Canadian cultivars shared the tNative haplotype, and the tEmmer haplotype was found in durum wheat introgressed from Td161 and a few durum wheat experimental lines from Austria (**Supplementary Table 4**). Allele effect analysis identified that the tEmmer type of 3B.2 conferred an effect that increased disease susceptibility (**Figure 6**).

There were a small number of QTL that did not co-locate with DTA or HT QTL. These include 1A.1, 1A.2, 6A.1, and 7A.3. The QTL 1A.1, located near the distal end of the short arm of chromosome 1A within a region from 13 to 20 Mb, was only significant for INC. Also on 1A was 1A.2, which mapped to 366 Mb on chromosome 1AL. It was detected in seven of the eight different environments, though not consistently across INC, IND and SEV, and a minor association with HT was also identified in one environment at this locus (**Figure 4**, **Table 2**, and **Supplementary Table 3**). 6A.1 was positioned at 12–23 Mb on 6AS. It had a significant effect on FHB, explaining up to 16% of phenotypic variation. This QTL also had a very minor effect for both HT and DTA with each only observed in a single environment. The 7A.3 QTL located to the distal region of 7A, around 671 Mb, had an effect on INC, SEV and IND, with no QTL for height or DTA found in this region (**Figure 4** and **Table 2**). This QTL was detected only in Brandon and Morden field sites, and explained up to 9.6% phenotypic variation (**Table 2**).

DISCUSSION

Phenotypic Data Analysis

The moderate to high heritability observed for FHB resistant traits in multiple environments in the durum AM panel indicated a large part of the phenotypic variation was contributed by genetic variation. The positive correlation between plant height and days to anthesis indicated that the genetic control of plant height and flowering time was partially shared (**Table 1**; Langer et al., 2014). The high proportion of disease susceptibility we observed in the field tests supports literature emphasizing the limited tetraploid wheat resources with a high level of FHB resistance (Oliver et al., 2008). The observed significantly negative correlations between FHB resistance and plant height and days to anthesis also agreed with previous findings summarized by Prat et al. (2014) and Steiner et al. (2017). Because the significant negative correlations between both DTA and HT and FHB traits ranged from -0.24^{***} to -0.60^{***} and -0.18^{*} to -0.42^{***} , respectively, there is considerable scope to shift this negative relationships (i.e., to have DTA more consistently around -0.24 and the correlation with HT toward -0.18). By



adopting strategies to stratify experimental genotypes into groups by both days to anthesis and plant height, it may be possible to recombine earlier to flower and shorter plants with reduced FHB symptoms. The correlation will not be broken but it can be shifted so that earlier maturing and shorter genotypes can be recombined with reduced FHB symptoms. Using this strategy, the negative relationship between plant height and FHB traits has been shifted by recombining semi-dwarf stature with a moderate level of resistance in hexaploid wheat cultivars such as Carberry (DePauw et al., 2011) and AAC Brandon (Cuthbert et al., 2017), both of which became widely adopted by producers. Adopting this strategy in durum wheat genetic enhancement could prove equally effective.

Genetic Architecture of FHB Resistance in the Durum AM Panel and Its Association With Flower Time and Plant Height

Compared to common wheat, durum wheat has limited genetic variation, and less effort has been committed to improve durum resistance to FHB (Buerstmayr et al., 2009, 2019; Prat et al., 2014, 2017). Within the current study, we identified a large number of QTL associated with FHB resistance with GWAS analysis from multiple environments and sources, broadening the resistance gene pool in durum wheat. The minor effect of these multiple QTL reinforces what is already known about the polygenic nature

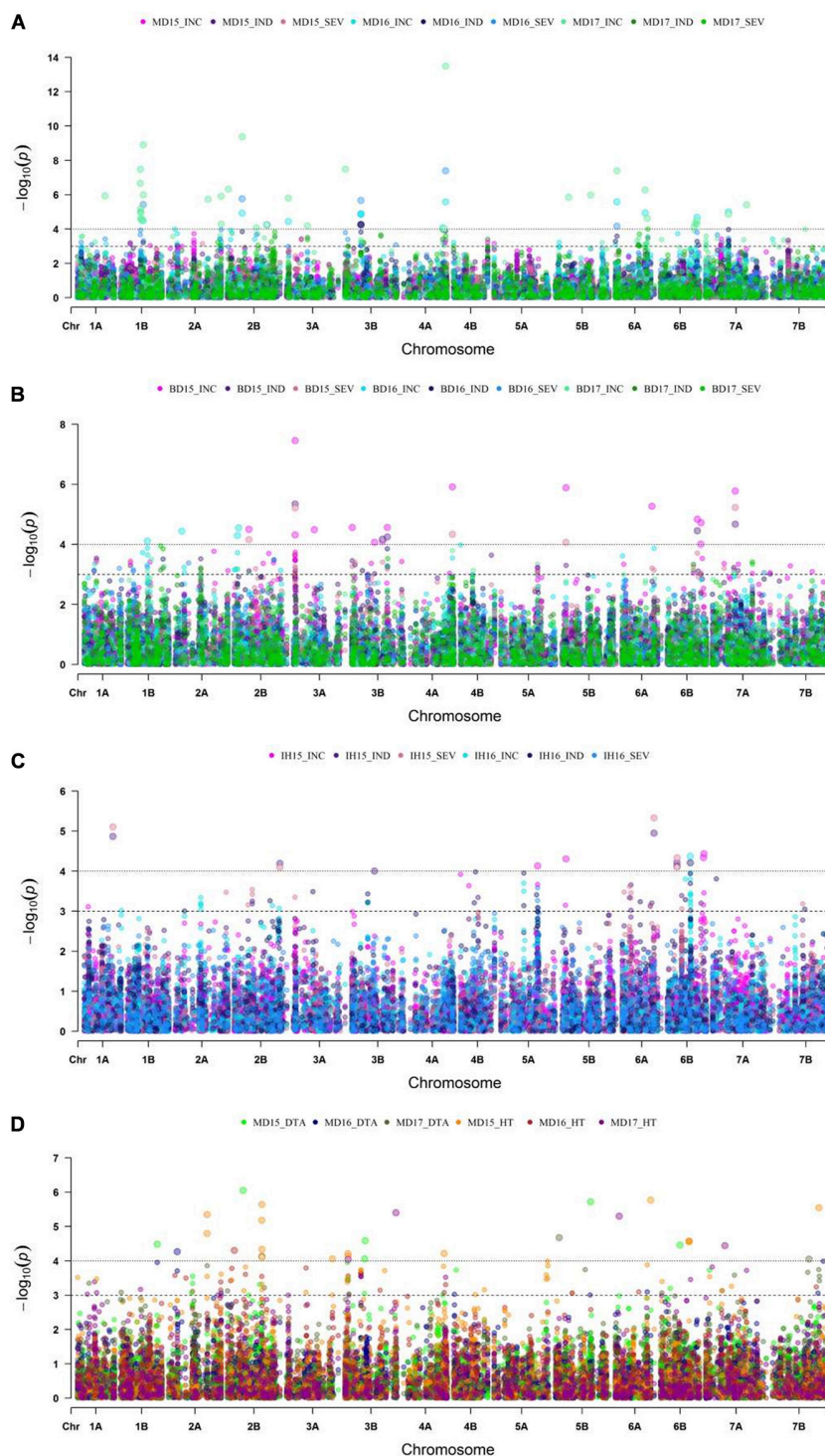


FIGURE 4 | Manhattan plots displaying genome wide marker-trait association analysis for FHB incidence (INC), index (IND) and severity (SEV) at **(A)** Morden, MB from the years 2015 to 2017; **(B)** Brandon, MB for years 2015 to 2017; **(C)** Indian Head, SK from 2015 to 2016 (with natural infection); and for **(D)** plant height (HT) and day to anthesis (DTA) at Morden, MB for 2015 to 2017 trials.

of FHB resistance, but also reveals the necessity of combining genes from multiple sources (Buerstmayr et al., 2009, 2019; Liu et al., 2009).

The major and most consistent FHB QTL found in previous studies is the hexaploid wheat Sumai 3 derived *Fhb1*, located on 3BS around 7.6–13.9 Mb (Anderson et al., 2001; Liu et al., 2006).

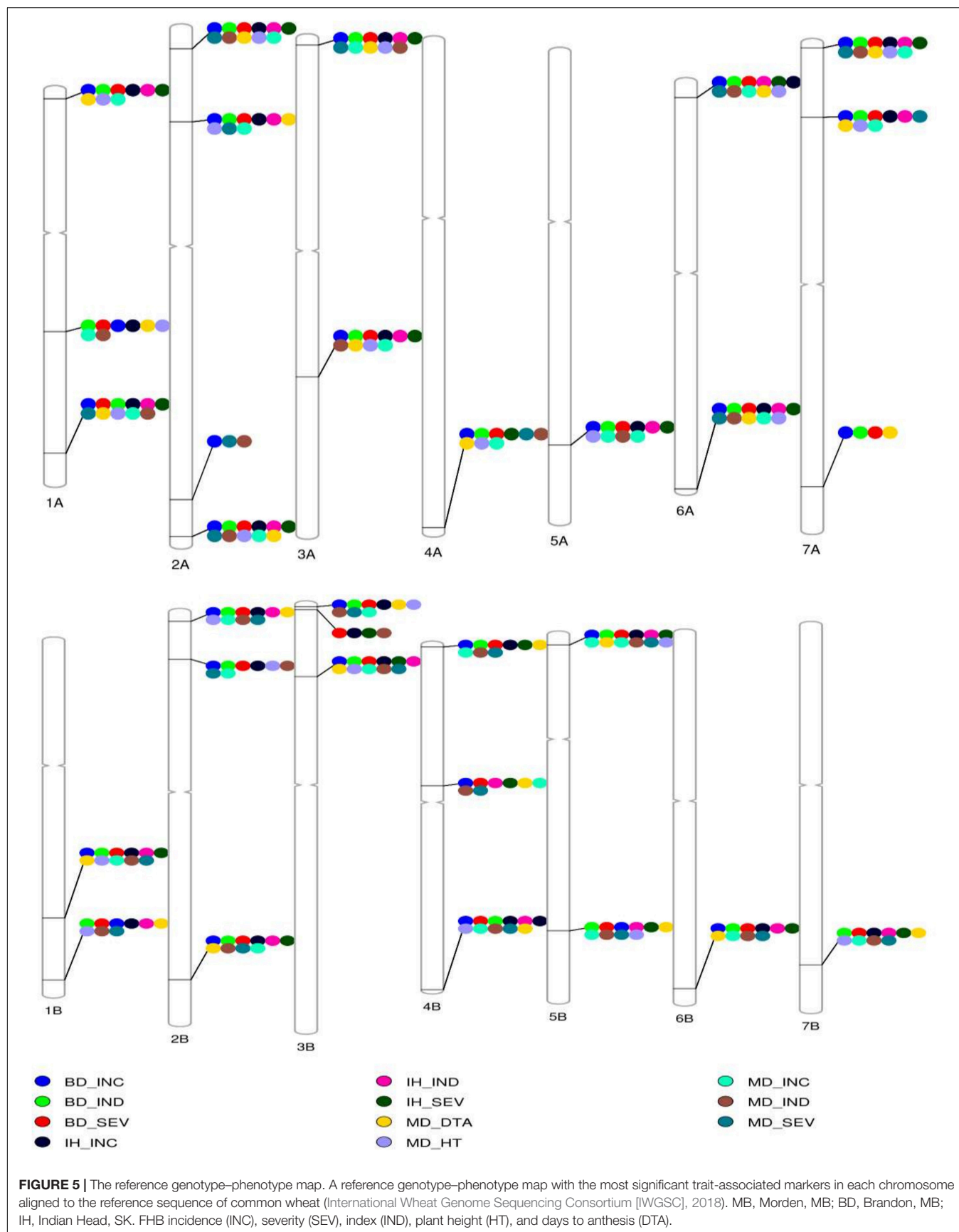


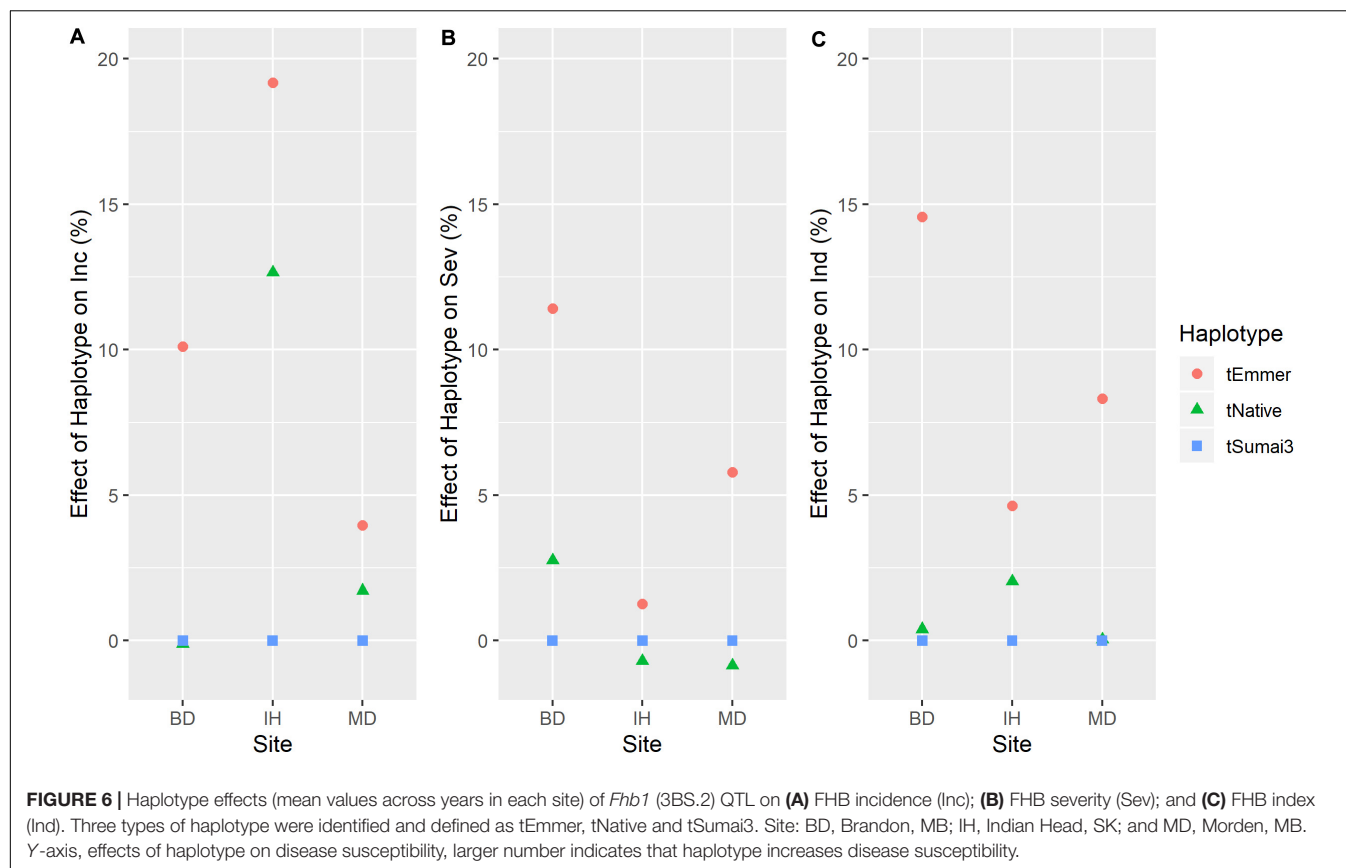
TABLE 2 | Quantitative trait loci names, physical positions, associated traits, explained phenotypic variance and significance of association with Fusarium head blight incidence (INC), index (IND), severity (SEV), days to anthesis (DTA), and plant height (HT) identified from durum association mapping panel across environments.

| QTL name | Physical position (Mb) | DTA | | | HT | | | FHB incidence | | | | | | | FHB index | | | | | | | FHB severity | | | | | | | Max R ² | | | |
|-------------|------------------------|--------|------------|------------|------------|------------|------------|---------------|------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|------------|------------|--------------|------------|------------|------------|------------|------------|------------|--------------------|------------|------------|--------|
| | | MD2015 | MD2016 | MD2017 | MD2015 | MD2016 | MD2017 | BD2015 | BD2016 | BD2017 | IH2015 | IH2016 | MD2015 | MD2016 | MD2017 | BD2015 | BD2016 | BD2017 | IH2015 | IH2016 | MD2015 | MD2016 | MD2017 | BD2015 | BD2016 | BD2017 | IH2015 | IH2016 | | MD2015 | MD2016 | MD2017 |
| 1A.1 | 13.4–20.9 | 0.3 | 1.2 | 0.3 | 0.8 | 0.2 | 0.5 | 0.9 | 1.8 | 1.8 | 2.4 | 0.0 | 0.0 | 3.2 | 2.9 | 0.8 | 1.5 | 0.2 | 1.0 | 0.8 | 0.3 | 2.7 | 0.8 | 1.0 | 2.3 | 0.1 | 1.4 | 0.2 | 0.4 | 3.4 | 0.9 | 8.1 |
| 1A.2 | 366 | 0.0 | 0.4 | 0.3 | 0.2 | 0.3 | 1.9 | 3.6 | 0.9 | 2.1 | 2.1 | 0.5 | 0.7 | 2.1 | 1.5 | 3.4 | 1.9 | 2.6 | 0.5 | 0.4 | 1.4 | 1.8 | 1.7 | 3.2 | 1.2 | 2.5 | 0.5 | 0.2 | 1.2 | 1.9 | 2.0 | 8.6 |
| 1A.3 | 503–580 | 2.3 | 1.9 | 2.6 | 1.3 | 1.8 | 0.4 | 0.2 | 3.2 | 1.4 | 0.2 | 1.7 | 2.2 | 1.6 | 1.1 | 0.8 | 3.0 | 1.8 | 0.4 | 0.6 | 1.6 | 1.9 | 1.8 | 0.7 | 3.4 | 1.5 | 0.2 | 0.6 | 0.4 | 1.7 | 1.4 | 7.8 |
| 1B.1 | 544–581 | 1.9 | 0.2 | 4.9 | 0.0 | 1.6 | 1.4 | 2.7 | 4.3 | 3.2 | 1.6 | 1.7 | 0.5 | 4.4 | 7.9 | 1.6 | 1.9 | 0.3 | 1.5 | 1.8 | 0.2 | 3.0 | 2.1 | 2.1 | 2.6 | 1.0 | 1.4 | 0.2 | 0.0 | 5.2 | 1.5 | 20.3 |
| 1B.2 | 662–668 | 1.0 | 0.3 | 0.0 | 2.0 | 1.6 | 0.6 | 0.6 | 1.9 | 1.8 | 0.4 | 1.2 | 0.3 | 1.0 | 1.5 | 0.8 | 3.6 | 3.4 | 1.2 | 1.2 | 2.1 | 1.3 | 2.2 | 0.9 | 1.8 | 4.0 | 1.1 | 0.2 | 2.4 | 1.0 | 1.8 | 8.0 |
| 2A.1 | 30–31 | 2.6 | 3.4 | 0.6 | 3.0 | 2.0 | 2.6 | 2.8 | 3.5 | 1.4 | 2.0 | 2.6 | 3.3 | 3.7 | 1.8 | 2.6 | 2.5 | 1.2 | 0.5 | 1.8 | 1.7 | 3.3 | 2.9 | 1.9 | 3.7 | 2.2 | 0.4 | 1.1 | 0.2 | 1.7 | 2.2 | 9.0 |
| 2A.2 | 138–142 | 1.9 | 2.3 | 3.0 | 0.7 | 0.4 | 1.0 | 2.2 | 2.3 | 3.1 | 2.6 | 3.4 | 3.9 | 1.9 | 0.8 | 3.6 | 2.1 | 3.3 | 2.0 | 1.6 | 3.4 | 1.7 | 2.5 | 3.0 | 1.9 | 2.3 | 0.8 | 0.6 | 1.6 | 0.7 | 2.2 | 10.1 |
| 2A.3 | 713–717 | 0.2 | 2.4 | 3.3 | 4.7 | 1.8 | 0.6 | 3.9 | 2.3 | 1.5 | 1.8 | 0.3 | 0.3 | 0.1 | 1.9 | 1.6 | 0.6 | 0.3 | 0.1 | 0.1 | 0.9 | 0.7 | 1.1 | 2.6 | 0.7 | 0.6 | 0.3 | 0.7 | 0.7 | 1.2 | 1.7 | 9.5 |
| 2A.4 | 762–769 | 1.8 | 2.2 | 0.7 | 1.5 | 4.0 | 2.1 | 2.7 | 2.4 | 1.8 | 2.6 | 2.6 | 0.3 | 3.4 | 5.2 | 2.3 | 1.7 | 2.5 | 1.3 | 2.4 | 0.6 | 3.5 | 3.3 | 2.6 | 2.1 | 2.0 | 1.3 | 0.7 | 0.2 | 3.8 | 3.2 | 12.9 |
| 2B.1 | 8.6–22 | 3.0 | 2.6 | 4.3 | 3.2 | 3.2 | 1.7 | 1.8 | 4.3 | 2.0 | 0.3 | 1.9 | 2.8 | 3.6 | 6.0 | 2.0 | 3.7 | 1.7 | 0.3 | 1.8 | 0.3 | 2.3 | 2.6 | 2.7 | 2.2 | 0.0 | 1.1 | 1.0 | 0.2 | 3.5 | 2.6 | 15.1 |
| 2B.2 | 92–102 | 0.0 | 0.8 | 0.8 | 0.7 | 1.7 | 1.6 | 3.3 | 2.1 | 2.9 | 0.7 | 0.9 | 0.2 | 4.7 | 8.1 | 2.2 | 1.2 | 0.4 | 0.6 | 1.0 | 0.7 | 3.5 | 1.8 | 3.2 | 2.2 | 1.0 | 1.2 | 0.3 | 0.7 | 5.5 | 1.5 | 20.8 |
| 2B.3 | 717–781 | 0.4 | 1.5 | 3.3 | 0.4 | 0.9 | 1.8 | 3.1 | 3.0 | 2.0 | 1.8 | 1.8 | 1.4 | 1.6 | 2.6 | 2.5 | 1.6 | 1.6 | 1.6 | 1.8 | 0.9 | 1.9 | 2.8 | 3.1 | 1.8 | 1.9 | 1.0 | 1.7 | 0.4 | 1.6 | 3.3 | 7.9 |
| 3A.1 | 9.6–13 | 1.6 | 0.3 | 2.9 | 0.6 | 1.7 | 2.2 | 6.2 | 1.5 | 1.7 | 2.2 | 1.4 | 0.5 | 4.6 | 5.4 | 4.6 | 1.6 | 3.3 | 1.5 | 1.7 | 0.4 | 2.9 | 2.0 | 4.6 | 2.1 | 2.5 | 1.7 | 1.3 | 2.0 | 3.2 | 2.4 | 15.4 |
| 3A.2 | 512–556 | 0.3 | 0.3 | 0.6 | 3.8 | 4.5 | 2.6 | 4.1 | 2.0 | 1.6 | 1.0 | 2.8 | 0.5 | 2.5 | 4.8 | 2.3 | 1.4 | 1.4 | 0.5 | 3.7 | 0.6 | 1.5 | 4.0 | 2.6 | 0.8 | 1.3 | 0.6 | 1.8 | 0.2 | 2.8 | 3.9 | 11.9 |
| 3B.1 | 3.7 | 1.4 | 0.8 | 4.0 | 0.9 | 1.4 | 2.3 | 4.1 | 3.2 | 2.9 | 2.5 | 0.8 | 0.1 | 3.1 | 8.1 | 3.4 | 1.1 | 1.8 | 0.8 | 0.5 | 0.2 | 3.2 | 2.7 | 3.5 | 1.1 | 1.4 | 0.6 | 0.5 | 0.3 | 4.5 | 2.6 | 20.8 |
| 3B.2 | 9.8 | 1.0 | 0.8 | 1.2 | 0.7 | 0.4 | 0.8 | 1.6 | 0.4 | 1.1 | 3.2 | 1.4 | 0.2 | 0.6 | 0.2 | 1.7 | 1.6 | 1.5 | 0.8 | 1.5 | 1.1 | 0.6 | 0.6 | 1.8 | 0.8 | 0.5 | 0.0 | 0.2 | 2.0 | 0.2 | 0.6 | 7.7 |
| 3B.3 | 148–233 | 0.0 | 1.4 | 2.2 | 4.5 | 3.5 | 2.9 | 3.0 | 2.3 | 1.7 | 2.1 | 2.5 | 0.1 | 3.8 | 1.8 | 2.5 | 2.9 | 1.5 | 1.4 | 3.2 | 0.1 | 3.7 | 2.4 | 2.6 | 2.6 | 0.8 | 1.8 | 2.3 | 0.0 | 2.5 | 2.3 | 9.2 |
| 4A.1 | 664–737 | 1.7 | 0.9 | 4.9 | 3.5 | 2.8 | 3.3 | 4.5 | 3.4 | 3.8 | 0.6 | 0.6 | 0.7 | 5.5 | 11.8 | 2.1 | 0.8 | 1.5 | 0.6 | 0.1 | 0.8 | 3.8 | 3.0 | 3.4 | 2.2 | 1.6 | 1.6 | 1.6 | 0.4 | 7.2 | 3.3 | 31.8 |
| 4B.1 | 3.8 | 0.9 | 1.5 | 0.4 | 0.7 | 1.0 | 1.3 | 2.5 | 4.0 | 0.6 | 4.3 | 0.3 | 2.0 | 3.2 | 2.0 | 2.4 | 1.9 | 0.3 | 2.3 | 0.3 | 1.0 | 2.8 | 3.3 | 1.7 | 1.5 | 0.3 | 1.3 | 0.5 | 0.6 | 1.8 | 3.4 | 10.5 |
| 4B.2 | 197–347 | 0.4 | 1.5 | 0.3 | 0.1 | 0.6 | 0.3 | 1.3 | 1.9 | 3.7 | 1.2 | 0.7 | 0.2 | 1.8 | 1.9 | 0.5 | 0.9 | 2.2 | 1.4 | 0.7 | 0.1 | 1.8 | 2.6 | 0.6 | 1.0 | 1.6 | 1.3 | 2.0 | 0.1 | 1.1 | 2.7 | 7.3 |
| 4B.3 | 673 | 0.0 | 0.4 | 2.0 | 2.7 | 0.6 | 1.8 | 1.5 | 2.5 | 2.3 | 2.2 | 3.1 | 3.1 | 2.0 | 3.0 | 2.3 | 1.8 | 2.4 | 2.2 | 3.0 | 3.1 | 2.6 | 4.2 | 2.4 | 2.5 | 2.7 | 1.4 | 3.2 | 2.2 | 1.8 | 3.8 | 8.8 |
| 5A.1 | 585–591 | 0.4 | 0.6 | 0.4 | 1.7 | 1.1 | 0.5 | 3.8 | 1.2 | 2.0 | 4.3 | 2.6 | 2.7 | 0.5 | 0.3 | 3.5 | 1.9 | 1.5 | 1.5 | 3.5 | 1.7 | 1.1 | 0.3 | 3.6 | 2.2 | 1.1 | 0.9 | 2.2 | 0.6 | 0.8 | 0.2 | 8.8 |
| 5B.1 | 19.5 | 0.2 | 1.5 | 4.6 | 1.0 | 0.9 | 2.0 | 6.6 | 2.8 | 2.1 | 5.2 | 0.8 | 1.0 | 1.1 | 2.6 | 4.2 | 1.0 | 0.6 | 2.5 | 0.9 | 1.0 | 1.3 | 1.4 | 5.0 | 0.7 | 0.1 | 1.4 | 1.6 | 0.4 | 1.3 | 1.5 | 14.5 |
| 5B.2 | 577–691 | 1.5 | 2.1 | 1.3 | 0.1 | 1.4 | 0.7 | 1.2 | 2.8 | 1.7 | 1.3 | 1.2 | 3.1 | 2.9 | 3.7 | 1.5 | 2.1 | 1.4 | 0.1 | 2.1 | 2.5 | 2.5 | 1.7 | 0.4 | 1.1 | 0.0 | 0.1 | 1.8 | 1.5 | 2.8 | 1.5 | 9.6 |
| 6A.1 | 12–23 | 0.7 | 1.3 | 1.4 | 0.3 | 1.5 | 1.3 | 2.7 | 2.4 | 3.2 | 0.6 | 2.4 | 0.6 | 5.8 | 6.3 | 2.4 | 2.6 | 2.4 | 0.2 | 1.4 | 0.5 | 3.8 | 3.4 | 2.1 | 3.3 | 2.9 | 0.5 | 1.4 | 2.2 | 4.0 | 3.4 | 15.7 |
| 6A.2 | 601–694 | 1.7 | 1.8 | 1.9 | 0.1 | 0.5 | 2.5 | 4.1 | 3.5 | 1.6 | 1.7 | 0.4 | 0.6 | 5.2 | 6.4 | 2.2 | 0.9 | 2.2 | 3.8 | 1.0 | 0.8 | 2.6 | 3.7 | 2.0 | 1.5 | 2.6 | 3.3 | 1.7 | 1.5 | 1.9 | 4.2 | 16.0 |
| 6B.1 | 585–707 | 1.8 | 1.8 | 1.9 | 0.5 | 0.2 | 0.4 | 4.0 | 3.0 | 1.4 | 2.2 | 0.0 | 0.3 | 5.0 | 4.1 | 2.5 | 1.0 | 0.2 | 2.5 | 0.2 | 0.1 | 3.5 | 1.1 | 2.5 | 0.5 | 0.5 | 1.1 | 0.1 | 0.1 | 4.0 | 0.9 | 12.2 |
| 7A.1 | 7.5–12 | 2.0 | 0.5 | 0.8 | 1.8 | 0.9 | 3.1 | 1.4 | 2.0 | 3.7 | 2.7 | 1.4 | 1.4 | 2.7 | 5.5 | 3.6 | 1.6 | 2.1 | 0.6 | 1.5 | 1.1 | 1.9 | 2.0 | 4.0 | 1.4 | 0.2 | 1.8 | 1.0 | 0.1 | 2.7 | 1.9 | 11.1 |
| 7A.2 | 102–113 | 1.5 | 2.1 | 2.6 | 1.1 | 0.3 | 1.4 | 5.1 | 3.3 | 3.7 | 2.7 | 0.7 | 0.5 | 5.4 | 5.2 | 4.7 | 1.6 | 0.5 | 0.9 | 0.5 | 0.0 | 4.5 | 2.8 | 5.0 | 0.5 | 0.0 | 0.3 | 0.4 | 0.0 | 3.8 | 3.4 | 13.5 |
| 7A.3 | 671 | 0.2 | 0.3 | 0.2 | 0.5 | 0.3 | 0.2 | 3.9 | 0.6 | 1.8 | 2.7 | 0.7 | 0.2 | 2.1 | 1.7 | 3.0 | 1.9 | 1.8 | 0.8 | 1.1 | 0.4 | 2.3 | 0.4 | 3.0 | 1.9 | 2.5 | 0.3 | 0.8 | 0.5 | 2.0 | 0.4 | 9.6 |
| 7B.1 | 610–658 | 1.7 | 0.1 | 6.0 | 0.4 | 1.6 | 1.1 | 5.2 | 4.0 | 2.5 | 3.6 | 0.9 | 1.5 | 1.9 | 4.7 | 4.0 | 2.0 | 2.6 | 1.4 | 1.0 | 1.7 | 1.9 | 1.8 | 4.0 | 0.7 | 2.9 | 2.1 | 0.1 | 1.1 | 1.7 | 1.6 | 12.9 |

The highest $-\log_{10}$ (p -value) of the markers from the QTL is given across all traits measured from the field trials performed in Brandon (BD), Morden (MD), and Indian Head (IH) in the years 2015, 2016 and 2017. $-\log_{10}$ (p) > 3 are in bold, values above the stringent Bonferroni significance threshold are underlined, and values below $-\log_{10}$ (p) = 1.3 are not shown.

Physical position (Mb): physical location of SNP markers in the QTL from Chinese Spring assembly (International Wheat Genome Sequencing Consortium [IWGSC], 2018).

Max R²: highest value for explained phenotypic variation for the marker across the traits and environments tested, expressed as%.



Introgression of *Fhb1* into durum wheat has been challenging, with one possible reason being the unstable expression in a durum genetic background (Zhao et al., 2018). Recently, Prat et al. (2017) successfully introgressed *Fhb1* into durum wheat, and some of those introgression lines are part of this AM panel. A QTL was found in the same region as *Fhb1* in this study, designated 3B.2. This QTL was detected in limited environments with a minor effect. QTL 3B.2 had three distinct haplotypes (Supplementary Table 4), and compared to haplotypes of Sumai 3 (tSumai 3) and Canadian cultivars (tNative), the haplotype from the experimental lines derived from emmer wheat Td161 (tEmmer) conferred disease susceptibility (Supplementary Table 4). This finding confirms previous findings that the *Fhb1* region from Td161 contributed to disease susceptibility when compared to the susceptible durum wheat Floradur (Buerstmayr et al., 2012). The resistance haplotype found in the GWAS study by Steiner et al. (2019b) corresponds to the tNative haplotype presented in this study. The tNative haplotype is the only haplotype found in the Canadian and American cultivars presented in both studies, while both the tNative and tEmmer haplotypes exist in durum wheat from Austria, CIMMYT, ICARDA, Italy and Morocco (Steiner et al., 2019b). Altogether, these findings indicate that one of the two non-Sumai 3 *Fhb1* region haplotypes found in tetraploid wheat contributed to disease susceptibility when compared to the other. Further characterizing the region with additional markers is needed to help resolve the source of the alleles

and further understand the effects of the three haplotypes identified in this study.

Two additional 3B QTL were found significantly associated with all of the traits, 3B.1 in the telomeric region of 3BS, and 3B.3 in the centromeric region of the short arm (3BSc). Recently, Wu et al. (2019) reported a QTL positioned at 2.0 Mb on the reference sequence from elite Chinese common wheat germplasm, almost the same region as the 3B.1 identified in this durum AM panel study. The 3B.3 QTL was one of the most stable QTL identified, with a larger effect on FHB resistance than other QTL in this AM panel. Notably, the resistant 3BSc haplotypes were identified in the durum wheat lines that also had *Fhb1* introgressed from Sumai 3 by Prat et al. (2017). The location of 3B.3 corresponds to the 3BSc region QTL previously reported as important to FHB resistance, particularly in Canadian elite germplasm, where 3BSc conferred a larger effect than *Fhb1* (McCartney et al., 2007). Also in agreement with findings from McCartney et al. (2007), the 3BSc QTL conferred a large effect on both plant height and DTA in elite Canadian wheat. Further research is needed to explore effects of *Fhb1*, 3B.1 and 3BSc in durum wheat.

QTL With No or Weak Association With Flowering Time and Height

The common association between plant height, flowering time and FHB resistance was illustrated in this study. Of the 31 FHB

QTL regions identified, all but five also had strong associations with plant height and/or flowering time. The relatively small effects of these QTL compared to other QTLs detected in this study may be related to the strong influence of flowering time on FHB resistance, potentially overinflating the effects of the QTL for FHB resistance due to the timing of flowering. Due to the progression of the FHB symptoms over time, the correlation between days to anthesis and disease development are confounded by the length of time for disease development. Due to cost constraints, disease rating was not evaluated over a time course to control for this effect, and thus we cannot exclude the observed correlation between FHB resistance and DTA may be caused by these confounding effects.

Fusarium head blight resistance QTL that are not associated with height or flowering time are much more appealing targets, as the negative influence of taller plants and complicated relationship with flowering time can be avoided. The targeted breeding of these QTL for resistance that do not carry extra undesirable traits will have the most likely success. The most favorable of these QTL may be 3B.2, but the QTL 1A.1, 1A.2, 6A.1, and 7A.3 with no association or weak association with DTA and HT are also desirable candidates. The 1A.1 QTL was located in the same region as the major QTL previously reported on the distal part of 1AS (summarized by Buerstmayr et al., 2009; Liu et al., 2009; Venske et al., 2019). Jiang et al. (2007a,b) located an FHB SEV QTL from the Chinese wheat line CJ9306 to position 27.2 Mb, and GWAS by Zhu et al. (2020) similarly identified an FHB QTL for IND from Chinese elite germplasm in the same region. A recent study by Sari et al. (2018) of *T. carthlicum* cv. Blackbird identified an important FHB QTL for INC, SEV and IND in the region of 1AS that agrees well with the 1A.1. The 1A.2 QTL colocalized with a QTL positioned at around 350 Mb for FHB severity and DON identified in Chinese elite germplasm (Wu et al., 2019) and for FHB resistance based on point inoculation in CIMMYT line C615 (Yi et al., 2018). In our study, we found this QTL was also associated with FHB incidence, index and severity. Within the AM panel of our study, although the resistance allele of 1A.1 was not found in Canadian cultivars, the 1A.2 occurred in several current Canadian cultivars with improved FHB resistance, including CDC Precision (Pozniak and Clarke, 2017b) and Brigade (Clarke et al., 2009; **Supplementary Table 3**).

The 6A.1 QTLs large effect on FHB resistance makes it appealing despite a small undesirable influence on DTA and HT. No major QTL clusters have been reported in a similar region as 6A.1, though Yi et al. (2018) reported a minor QTL in this region detected from a susceptible wheat line in one environment, and Lu et al. (2013) identified a minor QTL in the proximal 6A region for both FHB resistance and plant height. Because the 6A.1 resistance haplotype is present in a large number of Canadian durum wheat cultivars, including Brigade (Clarke et al., 2009), Transcend (Singh et al., 2012), CDC Credence (Sari et al., 2018) and CDC Precision (Pozniak and Clarke, 2017b; **Supplementary Table 3**), it should be possible for Canadian breeding programs to build on this resistance, though the effect of the QTL in Canadian elite durum cultivars remains to be validated.

The 7A.3 QTL, located at 671 Mb, with its relatively large effects on all FHB resistant traits without being associated with plant height or flowering time also make it another good target for breeding FHB resistance. Previous research identified a major QTL for type II resistance based on point inoculation in the vicinity of 7A.3 through the physical mapping of the SSRs *gwm276* and *gwm262* to positions of 642.9 and 681.4 Mb (Semagn et al., 2007; Buerstmayr et al., 2009). Wu et al. (2019) also reported a QTL affecting DON accumulation in the same region of elite Chinese germplasm, while Sari et al. (2018) reported QTL for SEV and IND in the same region from the durum wheat inbred line DT696.

From the durum AM panel in our study, 2A.2 located in the same region as a native durum FHB resistance QTL in previous research in cultivars Ben by Zhang et al. (2014) and Joppa by Zhao et al. (2018). In addition, the QTL 2A.2 was also found consistently associated with DTA, suggesting it plays a role in controlling flowering. In this durum AM panel, the resistance haplotype of 2A.2 was found in DT696 (Sari et al., 2018), an adapted source of FHB resistance in durum wheat, as well as several Canadian cultivars with improved FHB resistance derived from this line, including Brigade (Clarke et al., 2009), Transcend (Singh et al., 2012) CDC Credence (Sari et al., 2018), and CDC Precision (Pozniak and Clarke, 2017b; **Supplementary Table 3**). Despite its association with DTA, the effectiveness of the 2A.2 in native durum cultivars from Canada and United States make it another good target to breed durum wheat with improved FHB resistance.

QTL Co-located With Flowering Genes

The majority of the QTL identified from this AM panel were found associated with flowering time and/or plant height. As mentioned previously, the Notably, three QTL pairs, including 1A.3 and 1B.1, 2A.1 and 2B.1, and 5A.1 and 5B.2, were found in syntenic regions of the A/B genome that harbor known orthologous gene pairs controlling flower time. 1A.3 was in a similar region of a major QTL found in United States winter wheat cultivar NC-Neuse (Petersen et al., 2016, 2017). The *FLOWERING LOCUS T3-A1* (*TaFT3-A1*) gene that promotes flowering was found physically mapped around 528.1 Mb of 1A in CS Ref 1.0 (Zikhali et al., 2017; International Wheat Genome Sequencing Consortium [IWGSC], 2018), which is close to the region of 1A.3. The major QTL 1B.1 located to the region coinciding with a QTL of FHB resistance from the European winter wheat Arina (Semagn et al., 2007; Buerstmayr et al., 2009; Liu et al., 2009), as well as loci controlling DTA identified in the recent durum wheat GWAS by Steiner et al. (2019b). This QTL conferred a stable and large effect for INC, SEV, HT and DTA. Recently, the photoperiod gene *FLOWERING LOCUS T3-B1* (*TaFT3-B1*) that promotes flowering time, was physically identified at position 581 Mb of 1B (Zikhali et al., 2017), the same region as 1B.1. The 1B.1 and 1A.3 QTL occur in syntenic region of the genome, indicating the orthologous gene pair, *TaFT3-B1* and *TaFT3-A1*, as candidate genes underling the QTL effect in these regions.

The 2A.1 QTL conferred main effects for INC, IND, DTA and HT, physically positioned to around 27–31 Mb on chromosome

2A. This location is very near to the photoperiod gene *Ppd1A*, which has an important role in controlling flowering time and height, indicating 2A.1 as candidate gene controlling the QTL. Giancaspro et al. (2016) found a similar QTL positioned at 10 Mb on 2AS for FHB resistance in durum wheat, derived from the introgression of FHB resistance from Sumai 3, but with no report on its association with plant height. Gadaleta et al. (2019) identified a wall-associated receptor-like kinase (WAK2) in this region as the candidate gene for FHB resistance. Our study found a 2B QTL, designated 2B.1 that colocalizes with *Ppd1B* located in a syntenic region of 2A.1. This QTL contributed to INC, SEV, IND, DTA and HT. Thus, our findings support the *Ppd* loci on 2AS and 2BS as candidate genes responsible for the observed effects, although further studies with well stratified plant height and FHB rating DTA are required in order to explore the factors underlying these QTL.

Both the QTL 5A.1 on 5AL and 5B.2 on 5BL occur in syntenic regions that harbor orthologs of the well-known vernalization genes *VRNA1* (at 585.1 Mb) and *VRNB1* (at 613.0 Mb). 5A.1 and 5B.2 both conferred a stable effect for INC and IND, and while 5B.2 also had a large effect of on DTA, 5A.1 had no effect on DTA and only a minor effect on HT in one environment. Sari et al. (2018) reported a major FHB resistance QTL from the Canadian durum wheat line DT696 in the same region as 5A.1, also finding no DTA or HT QTL in this region. Xu et al. (2020) found QTL located in the same regions as 5A.1 and 5B.2 in common wheat that controlled anther extrusion, heading time and FHB resistance. There is potential that these vernalization genes are responsible for the FHB resistance coming from these regions, and that the *VRNA1* gene has just a minor effect on flowering time in durum wheat. The resistance haplotype of 5A.1 was found in Canadian durum cultivars including Brigade (Clarke et al., 2009) and CDC Alloy (Pozniak and Clarke, 2017a; **Supplementary Table 3**). Because of the presence of the resistant haplotype in current durum cultivars, and the minor effect on flowering time, we believe the *VRNA1* region QTL from this study and Sari et al. (2018) is a good target to improve FHB resistance in durum wheat. However, there is still need for further research to explore the mechanism of colocalization between the vernalization genes and FHB resistance and their effect on flowering in durum.

CONCLUSION

With genome wide association analysis we identified 31 QTL for FHB resistance. This confirms the quantitative nature and polygenic control of the FHB resistance and also signifies that this durum AM panel contains a large amount of genetic variation for FHB resistance loci. These QTL capture a large amount of the major QTL reported for hexaploid and tetraploid wheat which should facilitate improving FHB resistance in durum wheat. Five QTL found primarily for FHB resistance, including 1A.1, 1A.2, 5A.1, 6A.1, and 7A.3, could be used as initial targets to improve resistance in durum wheat without detrimental effects. Although 2A.2 is associated with DTA, the resistant haplotype exists in several Canadian and United States

cultivars with improved FHB resistance, and we think that due to its adaption to durum cultivars in North America it is also a good target. The majority of these QTL identified were associated with plant height and/or flowering time, indicating that phenology, flowering and height genes formed a complex network affecting FHB resistance in durum wheat. Prior knowledge of the haplotypes of these genes in breeding materials will provide an informed approach to stack these genes and give breeders the ability to design a better strategy to use these sources to improve FHB resistance. However, more research is needed to identify the mechanism of the trait associations, and truly determine whether pleiotropic effects of same gene, linkage drag of resistant genes, and/or disease escape due to flowering time and plant height are in effect. Only by completely understanding these relationships, can a better strategy, from genetic, genomics and breeding perspectives be developed to significantly increase FHB resistance in durum wheat. Finally, considering the attributes of QTL identified in this study, including the large number of minor effects, the varied expression across environments, and the complex interaction with flowering time and height, we suggest intercrossing the multiple sources of resistance. Then the progeny should be selected using a multi-trait based, high-throughput marker assisted selection approach that incorporates resistance, flowering time and height loci, in combination with intensive phenotyping, with the genotypes grouped by days to flower and plant height, across multiple target environments, as the most promising approach to develop durum wheat with a better level of resistance.

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

YR, RD, RK, PF, and WZ conceived and designed this study. YR and RK constructed the durum association mapping (AM) population. HB, RD, RK, RC, and YR contributed to the durum germplasm development in the AM population. YR, RK, HC, and SB contributed to seed increase of this AM population and the design of field trials. MH, AB, and SK conducted field trials and disease evaluation at FHB nurseries located in Morden and Brandon, Manitoba. YR, SB, RK, RC, and HC performed field trials and disease evaluation at FHB nursery in Indian Head, Saskatchewan. KB and BP contributed to the 90K SNP genotyping and marker identification. KB, WZ, and YR analyzed the data and interpreted results. KB, WZ, RR, and YR contributed to data validation. WZ, KB, and YR wrote the original draft. KB, RK, WZ, YR, RD, HB, HC, and PF reviewed and edited the manuscript. YR was the principal investigator and supervised the project. All authors contributed to the article and approved the submitted version.

FUNDING

Financial support was received from the Saskatchewan Agriculture Development Fund to YR and WZ, the SeCan to YR, the Agriculture and Agri-Food Canada and Western Grains Research Foundation to YR, RK, and RC, and the Canadian Wheat Improvement Flagship funded by the National Research Council Canada to PF.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of Saskatchewan Agriculture Development Fund, SeCan, Agriculture and Agri-food Canada (AAFC), Western Grain Research Foundation,

and National Research Council Canada (NRC) in carrying out this study. We thank Janet Condie at NRC for technical assistance with the 90K SNP genotyping, and Daoquan Xiang for critically reviewing the manuscript. The technical and field support at the Swift Current Research and Development Centre, Morden Research and Development Centre, and Brandon Research and Development Centre from AAFC is greatly appreciated.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2020.592064/full#supplementary-material>

REFERENCES

- Anderson, J. A., Stack, R., Liu, S., Waldron, B., Fjeld, A. D., Coyne, C., et al. (2001). DNA markers for Fusarium head blight resistance QTLs in two wheat populations. *Theor. Appl. Genet.* 102, 1164–1168. doi: 10.1007/s001220000509
- Bai, G., and Shaner, G. (1994). Scab of wheat: perspective and control. *Plant Dis.* 78, 760–766. doi: 10.1094/pd-78-0760
- Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to Fusarium head blight. *Annu. Rev. Phytopathol.* 42, 135–161. doi: 10.1146/annurev.phyto.42.040803.140340
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed effects models using lme4. *J. Stat. Softw.* 67, 1–48.
- 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
- Breseghele, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177. doi: 10.1534/genetics.105.044586
- Buerstmayr, H., Ban, T., and Anderson, J. A. (2009). QTL mapping and marker-assisted selection for Fusarium head blight resistance in wheat: a review. *Plant Breed.* 128, 1–26. doi: 10.1556/crc.36.2008.suppl.b.1
- Buerstmayr, H., Lemmens, M., Hartl, L., Doldi, L., Steiner, B., Stierschneider, M., et al. (2002). Molecular mapping of QTLs for Fusarium head blight resistance in spring wheat. I. Resistance to fungal spread (type II resistance). *Theor. Appl. Genet.* 104, 84–91. doi: 10.1007/s001220200009
- Buerstmayr, M., Huber, K., Heckmann, J., Steiner, B., Nelson, J. C., and Buerstmayr, H. (2012). Mapping of QTL for Fusarium head blight resistance and morphological and developmental traits in three backcross populations derived from *Triticum dicoccum* x *Triticum durum*. *Theor. Appl. Genet.* 125, 1751–1765. doi: 10.1007/s00122-012-1951-2
- Buerstmayr, M., and Buerstmayr, H. (2016). The semidwarfing alleles Rht-D1b and Rht-B1b show marked differences in their associations with anther-retention in wheat heads and with Fusarium head blight susceptibility. *Phytopathology* 106, 1544–1552. doi: 10.1094/phyto-05-16-0200-r
- Buerstmayr, M., Steiner, B., and Buerstmayr, H. (2019). Breeding for Fusarium head blight resistance in wheat — progress and challenges. *Plant Breed.* 139, 1–26. doi: 10.1111/pbr.12797
- Canadian Grain Commission (2017). Available online at: <https://www.grainscanada.gc.ca/en/grain-research/export-quality/cereals/wheat/western/annual-fusarium-damage/canada-western-amber-durum/>
- Chu, C., Niu, Z., Zhong, S., Chao, S., Friesen, T. L., Halley, S., et al. (2011). Identification and molecular mapping of two QTLs with major effects for resistance to Fusarium head blight in wheat. *Theor. Appl. Genet.* 123, 1107–1119. doi: 10.1007/s00122-011-1652-2
- Clarke, J. M., Knox, R. E., DePauw, R. M., Clarke, F. R., McCaig, T. N., Fernandez, M. R., et al. (2009). Brigade durum wheat. *Can. J. Plant Sci.* 89, 505–509.
- Cuthbert, P. A., Somers, D. J., and Brule-Babel, A. (2007). Mapping of Fhb2 on chromosome 6BS: a gene controlling Fusarium head blight field resistance in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 114, 429–437. doi: 10.1007/s00122-006-0439-3
- Cuthbert, R. D., DePauw, R. M., Knox, R. E., Singh, A. K., McCaig, T. N., McCallum, B., et al. (2017). AAC Brandon hard red spring wheat. *Can. J. Plant Sci.* 97, 393–401.
- DePauw, R. M., Knox, R. E., McCaig, T. N., Clarke, F. R., and Clarke, J. M. (2011). Carberry hard red spring wheat. *Can. J. Plant Sci.* 91, 529–534. doi: 10.4141/cjps10187
- 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
- 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
- Gadaleta, A., Colasuonno, P., Giove, S. L., Blanco, A., and Giancespro, A. (2019). Map-based cloning of *QFhb.mgb-2A* identifies a *WAK2* gene responsible for Fusarium head blight resistance in wheat. *Sci. Rep.* 9:6929.
- Gervais, L., Dedryver, F., Morlais, J. Y., Bodusseau, V., Negre, S., Bilous, M., et al. (2003). Mapping of quantitative trait loci for field resistance to Fusarium head blight in an European winter wheat. *Theor. Appl. Genet.* 106, 961–970. doi: 10.1007/s00122-002-1160-5
- Ghavami, F., Elias, E., Mamidi, S., Ansari, O., Sargolzaei, M., Adhikari, T., et al. (2011). Mixed model association mapping for Fusarium head blight resistance in Tunisian-derived durum wheat populations. *G3* 1, 209–218. doi: 10.1534/g3.111.000489
- Giancespro, A., Giove, S. L., Zito, D., Blanco, A., and Gadaleta, A. (2016). Mapping QTLs for Fusarium head blight resistance in an interspecific wheat population. *Front. Plant Sci.* 7:1381. doi: 10.3389/fpls.2016.01381
- Gilbert, J., and Haber, S. (2013). Overview of some recent research developments in Fusarium head blight of wheat. *Can. J. Plant Pathol.* 35, 149–174. doi: 10.1080/07060661.2013.772921
- Gilbert, J., and Tekauz, A. (2000). Review: recent developments in research on fusarium head blight of wheat in Canada. *Can. J. Plant Pathol.* 22, 1–8. doi: 10.1080/07060660009501155
- Haile, J. K., N'Diaye, A., Walkowiak, S., Nilsen, K. T., Clarke, J. M., Kutcher, H. R., et al. (2019). Fusarium head blight in durum wheat: recent status, breeding directions, and future research prospects. *Phytopathology* 109, 1664–1675. doi: 10.1094/phyto-03-19-0095-rvw
- Hao, Y., Rasheed, A., Zhu, Z., Wulff, B. B. H., and He, Z. (2019). Harnessing wheat Fhb1 for Fusarium resistance. *Trends Plant Sci.* 25, 1–3. doi: 10.1016/j.tplants.2019.10.006
- He, X., Dreisigacker, S., Singh, R. P., and Singh, P. K. (2019). Genetics for low correlation between Fusarium head blight disease and deoxynivalenol (DON) content in a bread wheat mapping population. *Theor. Appl. Genet.* 132, 2401–2411. doi: 10.1007/s00122-019-03362-9

- Huhn, M., Elias, E., Ghavami, F., Kianian, S. F., Chao, S., Zhong, S., et al. (2012). Tetraploid Tunisian wheat germplasm as a new source of Fusarium head blight resistance. *Crop Sci.* 52, 136–145. doi: 10.2135/cropsci2011.05.0263
- International Grains Council (2020). *Bulletin GMR 508*. Available online at: <https://www.igc.int/en/about/aboutus-meeting.aspx> (accessed July 13, 2020).
- International Wheat Genome Sequencing Consortium [IWGSC] (2018). Shifting the limits in wheat research, and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191
- Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233
- Jiang, G. L., Dong, Y., Shi, J., and Ward, R. W. (2007a). QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. II. Resistance to deoxynivalenol accumulation and grain yield loss. *Theor. Appl. Genet.* 115, 1043–1052. doi: 10.1007/s00122-007-0630-1
- Jiang, G. L., Shi, J., and Ward, R. W. (2007b). QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. I. Resistance to fungal spread. *Theor. Appl. Genet.* 116, 3–13. doi: 10.1007/s00122-007-0641-y
- Langer, S. M., Longin, C. F. H., and Würschum, T. (2014). Flowering time control in European winter wheat. *Front. Plant Sci.* 5:537. doi: 10.3389/fpls.2014.00537
- Li, G. Q., Zhou, J. Y., Jia, H. Y., Gao, Z. X., Fan, M., Luo, Y. J., et al. (2019). Mutation of a histidine-rich calcium-binding-protein gene in wheat confers resistance to Fusarium head blight. *Nat. Genet.* 51, 1106–1112. doi: 10.1038/s41588-019-0426-7
- 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
- Liu, S., Hall, M. D., Griffey, C. A., and McKendry, A. L. (2009). Meta-analysis of QTL associated with Fusarium head blight resistance in wheat. *Crop Sci.* 49, 1955–1968. doi: 10.2135/cropsci2009.03.0115
- Liu, S., Zhang, X., Pumphrey, M. O., Stack, R. W., Gill, B. S., and Anderson, J. A. (2006). Complex microcolinearity among wheat, rice, and barley revealed by fine mapping of the genomic region harboring a major QTL for resistance to Fusarium head blight in wheat. *Funct. Integr. Genomics* 6, 83–89. doi: 10.1007/s10142-005-0007-y
- Löffler, M., Schön, C. C., and Miedaner, T. (2009). Revealing the genetic architecture of FHB resistance in hexaploid wheat (*Triticum aestivum* L.) by QTL meta-analysis. *Mol. Breed.* 23, 473–488. doi: 10.1007/s11032-008-9250-y
- Lu, Q. X., Lillemo, M., Skinnies, H., He, X. Y., Shi, J. R., Ji, F., et al. (2013). Anther extrusion and plant height are associated with Type I resistance to Fusarium head blight in bread wheat line ‘Shanghai-3/Catbird’. *Theor. Appl. Genet.* 126, 317–334. doi: 10.1007/s00122-012-1981-9
- Lv, C., Song, Y., Gao, L., Yao, Q., Zhou, R., Xu, R., et al. (2014). Integration of QTL detection and marker assisted selection for improving resistance to Fusarium head blight and important agronomic traits in wheat. *Crop J.* 2, 70–78. doi: 10.1016/j.cj.2013.10.004
- McCartney, C. A., Somers, D. J., and Fedak, G. (2007). The evaluation of FHB resistance QTLs introgressed into elite Canadian spring wheat germplasm. *Mol. Breed.* 20, 209–221. doi: 10.1007/s11032-007-9084-z
- McMullen, M., Bergstrom, G., DeWolf, E., Dill-Macky, R., Hershman, D., Shaner, G., et al. (2012). A unified effort to fight an enemy of wheat and barley: fusarium head blight. *Plant Dis.* 96, 1712–1728. doi: 10.1094/pdis-03-12-0291-fe
- McMullen, M., Jones, R., and Gallenberg, D. (1997). Scab of wheat and barley: a re-emerging disease of devastating impact. *Plant Dis.* 81, 1340–1348. doi: 10.1094/pdis.1997.81.12.1340
- Mesterhazy, A. (1995). Types and components of resistance to Fusarium head blight of wheat. *Plant Breed.* 114, 377–386. doi: 10.1111/j.1439-0523.1995.tb00816.x
- Miedaner, T., and Korzun, V. (2012). Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* 102, 560–566. doi: 10.1094/phyto-05-11-0157
- Miller, J. D., Young, J. C., and Sampson, D. R. (1985). Deoxynivalenol and Fusarium head blight resistance in spring cereals. *J. Phytopathol.* 113, 359–367. doi: 10.1111/j.1439-0434.1985.tb04837.x
- Oliver, R. E., Cal, X., Friesen, T. L., Halley, S., Stack, R. W., and Xu, S. S. (2008). Evaluation of Fusarium head blight resistance in tetraploid wheat (*Triticum turgidum* L.). *Crop Sci.* 48, 213–222. doi: 10.2135/cropsci2007.03.0129
- Petersen, S., Lyerly, J. H., Maloney, P. V., Brown-Guedira, G., Cowger, C., Costa, J. M., et al. (2016). Mapping of head blight resistance quantitative trait loci in winter wheat cultivar NC-neuse. *Crop Sci.* 56, 1473–1483. doi: 10.2135/cropsci2015.05.0312
- Petersen, S., Lyerly, J. H., McKendry, A. L., Islam, M. S., Brown-Guedira, G., Cowger, C., et al. (2017). Validation of Fusarium head blight resistance QTL in US winter wheat. *Crop Sci.* 57, 1–12. doi: 10.2135/cropsci2015.07.0415
- Pozniak, C. J., and Clarke, J. M. (2017a). CDC Alloy durum wheat. *Can. J. Plant Sci.* 97, 385–389.
- Pozniak, C. J., and Clarke, J. M. (2017b). CDC Precision durum wheat. *Can. J. Plant Sci.* 97, 344–348.
- Prat, N., Buerstmayr, M., Steiner, B., Robert, O., and Buerstmayr, H. (2014). Current knowledge on resistance to Fusarium head blight in tetraploid wheat. *Mol. Breed.* 34, 1689–1699. doi: 10.1007/s11032-014-0184-2
- Prat, N., Guilbert, C., Prah, U., Wachter, E., Steiner, B., Langin, T., et al. (2017). QTL mapping of Fusarium head blight resistance in three related durum wheat populations. *Theor. Appl. Genet.* 130, 13–27. doi: 10.1007/s00122-016-2785-0
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- R Core Team (2017). *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ramasamy, R. K., Ramasamy, S., Bindroo, B. B., and Naik, V. G. (2014). STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus* 3:431. doi: 10.1186/2193-1801-3-431
- Rawat, N., Pumphrey, M. O., Liu, S., Zhang, X., Tiwari, V. K., Ando, K., et al. (2016). Wheat Fhb1 encodes a chimeric lectin with agglutinin domains and a pore-forming toxin-like domain conferring resistance to Fusarium head blight. *Nat. Genet.* 48, 1576–1580. doi: 10.1038/ng.3706
- 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
- Ruan, Y., Comeau, A., Langevin, F., Hucl, P., Clarke, J. M., Brule-Babel, A., et al. (2012). Identification of novel QTL for resistance to Fusarium head blight in a tetraploid wheat population. *Genome* 55, 853–864. doi: 10.1139/gen-2012-0110
- Sari, E., Berraies, S., Knox, R. E., Singh, A. K., Ruan, Y., Cuthbert, R. D., et al. (2018). High density genetic mapping of Fusarium head blight resistance QTL in tetraploid wheat. *PLoS One* 13:e0204362. doi: 10.1371/journal.pone.0204362
- Semagn, K., Skinnies, H., Bjornstad, A., Maroy, A. G., and Tarkegne, Y. (2007). Quantitative trait loci controlling Fusarium head blight resistance and low deoxynivalenol content in hexaploid wheat population from ‘Arina’ and NK93604. *Crop Sci.* 47, 294–303. doi: 10.2135/cropsci2006.02.0095
- Singh, A. K., Clarke, J. M., Knox, R. E., DePauw, R. M., McCaig, T. N., Fernandez, M. R., et al. (2012). Transcend durum wheat. *Can. J. Plant Sci.* 92, 809–813.
- Skinnies, H., Semagn, K., Tarkegne, Y., Maroy, A. G., and Bjornstad, A. (2010). The inheritance of anther extrusion in hexaploid wheat and its relationship to Fusarium head blight resistance and deoxynivalenol content. *Plant Breed.* 129, 149–155. doi: 10.1111/j.1439-0523.2009.01731.x
- Somers, D. J., Fedak, G., Clarke, J., and Cao, W. (2006). Mapping of FHB resistance QTLs in tetraploid wheat. *Genome* 49, 1586–1593. doi: 10.1139/g06-127
- Somers, D. J., Fedak, G., and Savard, M. (2003). Molecular mapping of novel genes controlling Fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. *Genome* 46, 555–564. doi: 10.1139/g03-033
- Srinivasachary, G. N., Steed, A., Hollins, T., Bayles, R., Jennings, P., and Nicholson, P. (2009). Semi-dwarfing Rht-B1 and Rht-D1 loci of wheat differ significantly in their influence on resistance to Fusarium head blight. *Theor. Appl. Genet.* 118, 695–702. doi: 10.1007/s00122-008-0930-0
- Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., and Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Trop. Plant Pathol.* 42, 165–174. doi: 10.1007/s40858-017-0127-7

- Steiner, B., Buerstmayr, M., Wagner, C., Danler, A., Eshonkulov, B., Ehn, M., et al. (2019a). Fine-mapping of the Fusarium head blight resistance QTL *Qfhs.ifa-5A* identifies two resistance QTL associated with anther extrusion. *Theor. Appl. Genet.* 132, 2039–2053. doi: 10.1007/s00122-019-03336-x
- Steiner, B., Michel, S., Maccaferri, M., Lemmens, M., Tuberosa, R., and Buerstmayr, H. (2019b). Exploring and exploiting the genetic variation of Fusarium head blight resistance for genomic-assisted breeding in the elite durum wheat gene pool. *Theor. Appl. Genet.* 132, 969–988. doi: 10.1007/s00122-018-3253-9
- Su, Z. Q., Bernardo, A., Tian, B., Chen, H., Wang, S., Ma, H. X., et al. (2019). A deletion mutation in TaHRC confers Fhb1 resistance to Fusarium head blight in wheat. *Nat. Genet.* 51, 1099–1105. doi: 10.1038/s41588-019-0425-8
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9, 1–9.
- Van Sanford, D., Anderson, J., Campbell, K., Costa, J., Cregan, P., Griffey, C., et al. (2001). Discovery and deployment of molecular markers linked to fusarium head blight resistance: an integrated system for wheat and barley. *Crop Sci.* 41, 638–644. doi: 10.2135/cropsci2001.413638x
- Venske, E., Dos Santos, R. S., Farias, D. D. R., Rother, V., Maia, L. C. D., Pegoraro, C., et al. (2019). Meta-analysis of the QTLome of Fusarium head blight resistance in bread wheat: refining the current puzzle. *Front. Plant Sci.* 10:727. doi: 10.3389/fpls.2019.00727
- Wang, R., Chen, J., Anderson, J. A., Zhang, J., Zhao, W., Wheeler, J., et al. (2017). Genome-wide association mapping of Fusarium head blight resistance in spring wheat lines developed in the Pacific Northwest and CIMMYT. *Phytopathology* 107, 1486–1495. doi: 10.1094/phyto-02-17-0073-r
- Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genome diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183
- Wei, T., and Simko, V. (2017). *Package 'Corrplot'*. Available online at: <https://CRAN.R-project.org/package=corrplot> (accessed October 16, 2017).
- Wu, L., Zhang, Y., He, Y., Jiang, P., Zhang, X., and Ma, H. X. (2019). Genome-wide association mapping of resistance to fusarium head blight spread and deoxynivalenol accumulation in Chinese elite wheat germplasm. *Phytopathology* 109, 1208–1216. doi: 10.1094/phyto-12-18-0484-r
- Xu, K. J., He, X. Y., Dreisigacker, S., He, Z. H., and Singh, P. K. (2020). Anther extrusion and its association with Fusarium head blight in CIMMYT wheat germplasm. *Agronomy* 10:47. doi: 10.3390/agronomy10010047
- Yi, X., Cheng, J., Jiang, Z., Hu, W., Bie, T., Gao, D., et al. (2018). Genetic analysis of Fusarium head blight resistance in CIMMYT bread wheat line C615 using traditional and conditional QTL mapping. *Front. Plant Sci.* 9:573. doi: 10.3389/fpls.2018.00573
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., and Yang, T. L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 5, 1786–1788. doi: 10.1093/bioinformatics/bty875
- Zhang, Q., Axtman, J. E., Faris, J. D., Chao, S., Zhang, Z., Friesen, T. L., et al. (2014). Identification and molecular mapping of quantitative trait loci for Fusarium head blight resistance in emmer and durum wheat using a single nucleotide polymorphism-based linkage map. *Mol. Breed.* 34, 1677–1687. doi: 10.1007/s11032-014-0180-6
- Zhao, M., Leng, Y., Chao, S., Xu, S. S., and Zhong, S. (2018). Molecular mapping of QTL for Fusarium head blight resistance introgressed into durum wheat. *Theor. Appl. Genet.* 131, 1939–1951. doi: 10.1007/s00122-018-3124-4
- Zhu, Z. W., Chen, L., Zhang, W., Yang, L. J., Zhu, W. W., Li, J. H., et al. (2020). Genome-Wide association analysis of Fusarium head blight resistance in Chinese elite wheat lines. *Front. Plant Sci.* 11:206. doi: 10.3389/fpls.2020.00206
- Zikhali, M., Wingen, L. U., Leverington-Waite, M., Specel, S., and Griffiths, S. (2017). The identification of new candidate genes *Triticum aestivum* FLOWERING LOCUS T3-B1 (TaFT3-B1) and TARGET OF EAT1 (TaTOE1-B1) controlling the short-day photoperiod response in bread wheat. *Plant Cell Environ.* 40, 2678–2690. doi: 10.1111/pce.13018

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 Ruan, Zhang, Knox, Berraies, Campbell, Ragupathy, Boyle, Polley, Henriquez, Burt, Kumar, Cuthbert, Fobert, Buerstmayr and DePauw. 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.



Genomic Breeding for Diameter Growth and Tolerance to *Leptocybe* Gall Wasp and *Botryosphaeria/Teratosphaeria* Fungal Disease Complex in *Eucalyptus grandis*

Makobatjati M. Mphahlele^{1,2}, Fikret Isik³, Gary R. Hodge^{3,4†} and Alexander A. Myburg^{2*†}

¹ Mondi Forests, Research and Development Department, Trahar Technology Centre – TTC, Hilton, South Africa,

² Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa, ³ Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States, ⁴ Camcore, North Carolina State University, Raleigh, NC, United States

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Jaroslav Klápšte,
New Zealand Forest Research
Institute Limited (Scion), New Zealand
Shogo Tsuruta,
University of Georgia, United States

*Correspondence:

Alexander A. Myburg
zander.myburg@up.ac.za

[†]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: 08 December 2020

Accepted: 29 January 2021

Published: 26 February 2021

Citation:

Mphahlele MM, Isik F, Hodge GR
and Myburg AA (2021) Genomic
Breeding for Diameter Growth
and Tolerance to *Leptocybe* Gall
Wasp
and *Botryosphaeria/Teratosphaeria*
Fungal Disease Complex
in *Eucalyptus grandis*.
Front. Plant Sci. 12:638969.
doi: 10.3389/fpls.2021.638969

Eucalyptus grandis is one of the most important species for hardwood plantation forestry around the world. At present, its commercial deployment is in decline because of pests and pathogens such as *Leptocybe invasa* gall wasp (*Lepto*), and often co-occurring fungal stem diseases such as *Botryosphaeria dothidea* and *Teratosphaeria zuluensis* (*BotryoTera*). This study analyzed *Lepto*, *BotryoTera*, and stem diameter growth in an *E. grandis* multi-environmental, genetic trial. The study was established in three subtropical environments. Diameter growth and *BotryoTera* incidence scores were assessed on 3,334 trees, and *Lepto* incidence was assessed on 4,463 trees from 95 half-sib families. Using the *Eucalyptus* EUChip60K SNP chip, a subset of 964 trees from 93 half-sib families were genotyped with 14,347 informative SNP markers. We employed single-step genomic BLUP (ssGBLUP) to estimate genetic parameters in the genetic trial. Diameter and *Lepto* tolerance showed a positive genetic correlation (0.78), while *BotryoTera* tolerance had a negative genetic correlation with diameter growth (−0.38). The expected genetic gains for diameter growth and *Lepto* and *BotryoTera* tolerance were 12.4, 10, and −3.4%, respectively. We propose a genomic selection breeding strategy for *E. grandis* that addresses some of the present population structure problems.

Keywords: ssGBLUP, genetic correlation, *Eucalyptus grandis*, *Leptocybe invasa*, *Botryosphaeria dothidea*, *Teratosphaeria zuluensis*

INTRODUCTION

Fast-growing plantation forests are essential to the pulp, paper, and timber industries and the emerging biorefinery and biomaterials industries (Perlack et al., 2005; Cetinkol et al., 2012; Devappa et al., 2015; Stafford et al., 2020). The sustainability of many of these industries is dependent on woody biomass from plantation-grown *Eucalyptus* trees. *Eucalyptus* species are adaptable,

fast-growing, generally resilient to pests and pathogens, and have the desired wood qualities for diverse wood products (Malan, 1993; Stafford et al., 2020). Volume growth and wood density are essential measures for forest plantation productivity (Raymond, 2002). However, pest and pathogen challenges have increased in severity in the past decades, posing a significant risk to *Eucalyptus* plantation forestry productivity and sustainability in subtropical regions (Wingfield et al., 2015). How to ensure continued genetic gains for volume growth in the presence of severe pest and pathogen challenges has become an essential question for plantation species such as *Eucalyptus grandis*.

Leptocybe invasa Fisher & La Salle is one of the most damaging insect pests of *Eucalyptus* species that affects growth by forming galls on leaves and leaf petioles. The insect is native to Queensland, Australia, known as the Blue Gum Chalcid wasp (Hymenoptera: Eupholidea). It has spread across the globe, infesting a wide range of commercially grown *Eucalyptus* species and their hybrids, resulting in severe losses in young plantations and nursery seedlings (Mendel et al., 2004; Nyeko et al., 2010; Chang et al., 2012; da Silva et al., 2020). First reported in the Mediterranean Basin and the Middle East in 2000 (Viggiani et al., 2000; Mendel et al., 2004), *L. invasa* subsequently spread throughout countries in Africa, America, and Asia (Nyeko, 2005; Wiley and Skelly, 2008; Prabhu, 2010; Zhu et al., 2012). Two parasitoid species of *L. invasa* from Australia, *Quadrastichus mendeli* and *Selitrichodes kryceri*, were deployed as biological controls to manage severe infestation levels in *Eucalyptus* plantations in Israel (Kim et al., 2008). Tracking the introduction of *L. invasa* in South Africa, *Q. mendeli* was recently discovered, and the biological control potential of *L. invasa* in South African *Eucalyptus* plantations was investigated (Bush et al., 2018). Another recently discovered parasitoid species of *L. invasa* from Australia, *S. neseri*, was described and investigated for its parasitism rates in South Africa, ranging from 9.7 to 71.8% (Dittrich-Schroder et al., 2014).

Resistance-linked DNA markers for molecular breeding is an alternative strategy to manage pest challenges. Towards this, simple sequence repeat (SSR) markers have been identified that jointly explained 3–37% of the variation of resistance in *E. grandis* and were validated in *E. tereticornis* explaining 24–48% of the variation of resistance (Zhang et al., 2018). Due to the significant variation that exists within and between *Eucalyptus* species, there is opportunity to breed for *L. invasa* tolerance (Mendel et al., 2004; Thu et al., 2009; Durand et al., 2011; Sangtongpraow et al., 2011; Dittrich-Schroder et al., 2012; Nugnes et al., 2015; Zheng et al., 2016). A recent genome-wide association study in an *E. grandis* breeding population identified candidate genomic regions on chromosomes 3, 7, and 8 that contained putative candidate genes for tolerance. These candidate genomic regions explained ~17.6% of the total phenotypic variation of *L. invasa* tolerance (Mhoswa et al., 2020).

Teratosphaeria zuluensis, a fungal pathogen that causes stem canker, previously known as *Coniothyrium* canker, is a devastating stem disease of *Eucalyptus* species and is one of the most severe pathogens of plantation-grown *Eucalyptus* spp. (Wingfield et al., 1996; Crous et al., 2009; Aylward et al., 2019). It was first recognized in South Africa in 1989 and described

in 1996 (Wingfield et al., 1996). *T. zuluensis* has been reported on *Eucalyptus* spp. in Malawi, Mozambique and Zambia (Jimu et al., 2015), Hawaii (Cortinas et al., 2004), Ethiopia (Gezahgne et al., 2003), and Argentina and Vietnam (Gezahgne et al., 2004b). Infections from *T. zuluensis* results in necrotic spots on green branches and the main stem, giving a “cat-eye” appearance that develops into large cankers on susceptible trees. *T. zuluensis* infection reduces wood quality by penetrating the cambium to form black kino filled pockets and may lead to tree death (Wingfield et al., 1996; Gezahgne et al., 2003).

Botryosphaeria dothidea is also a devastating fungal pathogen of eucalypt species affecting the stem. *B. dothidea* is known to have endophytic characteristics with instances of opportunistic latent infections (Smith et al., 1996; Slippers et al., 2009). Species of the *Botryosphaeriaceae* family infect plants via natural apertures (Bihon et al., 2011) and wounding (Epstein et al., 2008). *B. dothidea* infection results in longitudinal cracks that penetrate the bark into the xylem forming kino pockets in the wood, and stem cankers and tip dieback (Smith et al., 1994). It infects eucalypts in many countries including the Congo (Roux et al., 2000), Australia (Burgess et al., 2019), South Africa (Smith et al., 1994), Ethiopia (Gezahgne et al., 2004a), Venezuela (Mohali et al., 2007), Colombia (Rodas et al., 2009), Uruguay (Perez et al., 2008), and China (Chen et al., 2011). Field assessment of the two fungal stem pathogens has revealed that the symptoms of *B. dothidea* and *T. zuluensis* can be present separately or concurrently at varying levels on trees in the population in the form of a fungal stem disease complex.

In general, tree breeding strategies use pedigree information to estimate genetic merit, often in trials with large numbers of individuals in open-pollinated families. The availability of a reference genome sequence of *E. grandis* (Myburg et al., 2014) and the development of a robust single-nucleotide polymorphism (SNP, EUChip60K) chip for high-throughput genotyping in multiple eucalypt species (Silva-Junior et al., 2015) have created opportunities for implementing new breeding strategies based on the genomic prediction of breeding values. While conventional pedigree relationships represent the average proportion of shared alleles, SNP markers can track Mendelian segregation patterns enabling the detection of unknown (cryptic) relationships and more precise estimation of known relationships (Habier et al., 2007; Hayes et al., 2009; Hill and Weir, 2010). However, the genotyping of all individuals in large open-pollinated tree breeding populations would be prohibitively expensive. Single-step genomic (ssG)BLUP analysis is an attractive alternative that blends the known pedigree of the entire population with the genomic relationship matrix of a subset of genotyped individuals (Legarra et al., 2009; Misztal et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010). Thereby, ssGBLUP analysis extends the benefits of applying of genomic selection to non-genotyped individuals (Legarra et al., 2014), therefore allowing for multivariate and univariate analysis (Guo et al., 2014) in livestock (Lourenco et al., 2015; Ma et al., 2015) and forest trees (Ratcliffe et al., 2017; Klapste et al., 2018, 2020; Cappa et al., 2019).

Improving forest plantation productivity requires recurrent selection of multiple traits, such as growth, wood quality, and tolerance to pests and pathogens. A multivariate analysis involves

estimating genetic correlations between traits to understand their correlated responses (Burdon, 1977). The correlated phenotypes of growth and pest and disease traits are attributable to shared genetic factors (pleiotropy) and/or linked genetic factors (linkage disequilibrium) and their interrelationships with environment factors (Falconer and Mackay, 1996). Being able to partition these components will help improve breeding strategies for correlated traits (Chen and Lubberstedt, 2010).

In this study, we measured breeding trials of *E. grandis* composed of trees from three half-sib pedigree linked generations and some unrelated families for diameter growth at breast height, tolerance to stem disease caused by the co-occurrence of *B. dothidea* and *T. zuluensis* (*Botryotera*), and tolerance to leaf gall caused by *L. invasa* (*Lepto*). The study aimed to obtain genetic parameters and genetic gains for growth, pest, and pathogen tolerance in this multi-generation breeding trial comparing ABLUP (pedigree-based BLUP analysis) and ssGBLUP models. We further investigated the additive genetic correlations and genotype-by-environmental ($G \times E$) interactions of diameter growth and tolerance to *Lepto* and *Botryotera*. Based on the results, we discuss the utility of genomic selection in *E. grandis* for simultaneous improvement of growth and tolerance to the gall wasp and fungal stem disease.

MATERIALS AND METHODS

Breeding History and Phenotyping of the Study Population

Eucalyptus grandis W. Hill ex Maiden was introduced to South Africa in the early 1900s and included various government breeding populations as a timber resource for the mining industry. Private breeding programs only started in the early 1970s, initiated from government landrace breeding populations. Breeding objectives for these landrace breeding populations gradually shifted to target traits for pulp and paper products rather than timber production in successive generations and trial series (Figure 1). We had access to seed from two first-generation selections from the 2nd trial series in this study population, with 32 selections from the 3rd trial series as our third-generation families and 28 selections from the 4th trial series as our fourth-generation families (Supplementary Table 1). Also included in the study was 33 unrelated (no pedigree link) families as controls, with seed sourced in the early 1990s from selections in Swaziland. The 93 half-sib pedigree linked families and the 33 unrelated control families were planted across three sites Mtunzini, Kwambonambi, and Nyalazi in KwaZulu Natal, a sub-tropical region in South Africa (Figure 2 and Supplementary Table 1). Families from the different generations were planted together in the three trial sites. The experimental design was a randomized complete block planted at single tree plots at 15 replicates per family. Field tolerance to *Lepto* was assessed at age 1 using a four-scale incidence score in which trees with a score of 4 shows no evidence of an attack on the leaf midrib or petiole, a score of 3 shows evidence of an attack on the leaf midrib or petiole without galls, and a score of 2 indicates trees with an attack on the leaf midrib or petiole with galls. Trees with a score of 1 present a

lethal outcome from an attack on the leaf midrib or petiole with galls (Figure 3). Field tolerance to *Botryotera* was assessed at age 3 using an incidence score in which a score of 6 represents trees with no spots/cracks or redness and trees with a score 5 show symptoms of *T. zuluensis* spots with redness, whereas trees with a score of 4 have *B. dothidea* cracks with redness. Trees with a score of 3 shows symptoms with *T. zuluensis* spots and *B. dothidea* cracks with redness, and a score of 2 represents trees with heavy *T. zuluensis* spots, and *B. dothidea* cracks with redness, and a score of 1 represents trees with heavy *T. zuluensis* spots and *B. dothidea* cracks with redness and cankers (Figure 4). Diameter growth at breast height (1.3 m over-bark) was measured at age 4.

Genotyping of the Study Population

DNA was extracted from leaves using the NucleoSpin DNA extraction kit (Machery-Nagel, Germany). The *Eucalyptus* (EUChip60K) SNP chip as described by Silva-Junior (Silva-Junior et al., 2015) available from GeneSeek (Neogen, Lansing, MI, United States) was used to genotype 964 trees across the families and trials (Supplementary Table 1). Of the 95 families in the trials, 93 contained a subset of 964 genotyped trees ranging from 2 to 24 trees per family. The two second-generation families were not genotyped. An average of four trees per family were genotyped of the unrelated families. For the third generation, 15 trees per family were genotyped, while in the fourth generation, 14 per family were genotyped. Of the 64,639 markers on the SNP chip (Silva-Junior et al., 2015), there are a total of 14,347 informative SNP markers with GenTrain score ranging from 0.37 to 0.93. Retained markers had a call rate of above 90% and a minor allele frequency above 0.05. The SNP genotypes frequencies of the 14,347 markers were AA (0.307), GG (0.283), AG (0.270), CC (0.068), AC (0.065), and 0.007 missing. The number of SNP markers distributed on linkage groups ranged from 1018 (Chromosome 1) to 1877 (Chromosome 10). The SNP marker frequencies and distribution analysis were performed with the *synbreed* 0.10-2 R package (Wimmer et al., 2012) and the imputing of the missing SNP data based on allelic distribution, assuming Hardy-Weinberg equilibrium.

Statistical Analyses

Mixed Model Analysis

Linear mixed models were fit to estimate variance components and solve mixed model equations to obtain solutions for fixed and random effects. The matrix notation for the linear mixed models used is as follows:

$$y = X\beta + Zu + \varepsilon \quad (1)$$

where y is a vector of phenotypes, X is the design matrix for the fixed effects (site), β is the vector of the fixed effect coefficients (intercept site), Z is an incidence matrix for the random effects of individual trees, u is the vector of random effect coefficients (genotype, genotype by site interaction, replication effect nested in site effect), and ε is the vector of residual effect coefficients. The expectations of y , u , and ε are $E(y) = X\beta$, $E(u) = N(0, \sigma_u^2)$, and $E(\varepsilon) = 0$ and the variances are $Var(y) = V = ZGZ' + R$, $Var(\varepsilon) = R = N(0, I\sigma_e^2)$, and $Var(u) = G = A\sigma_u^2$, respectively,

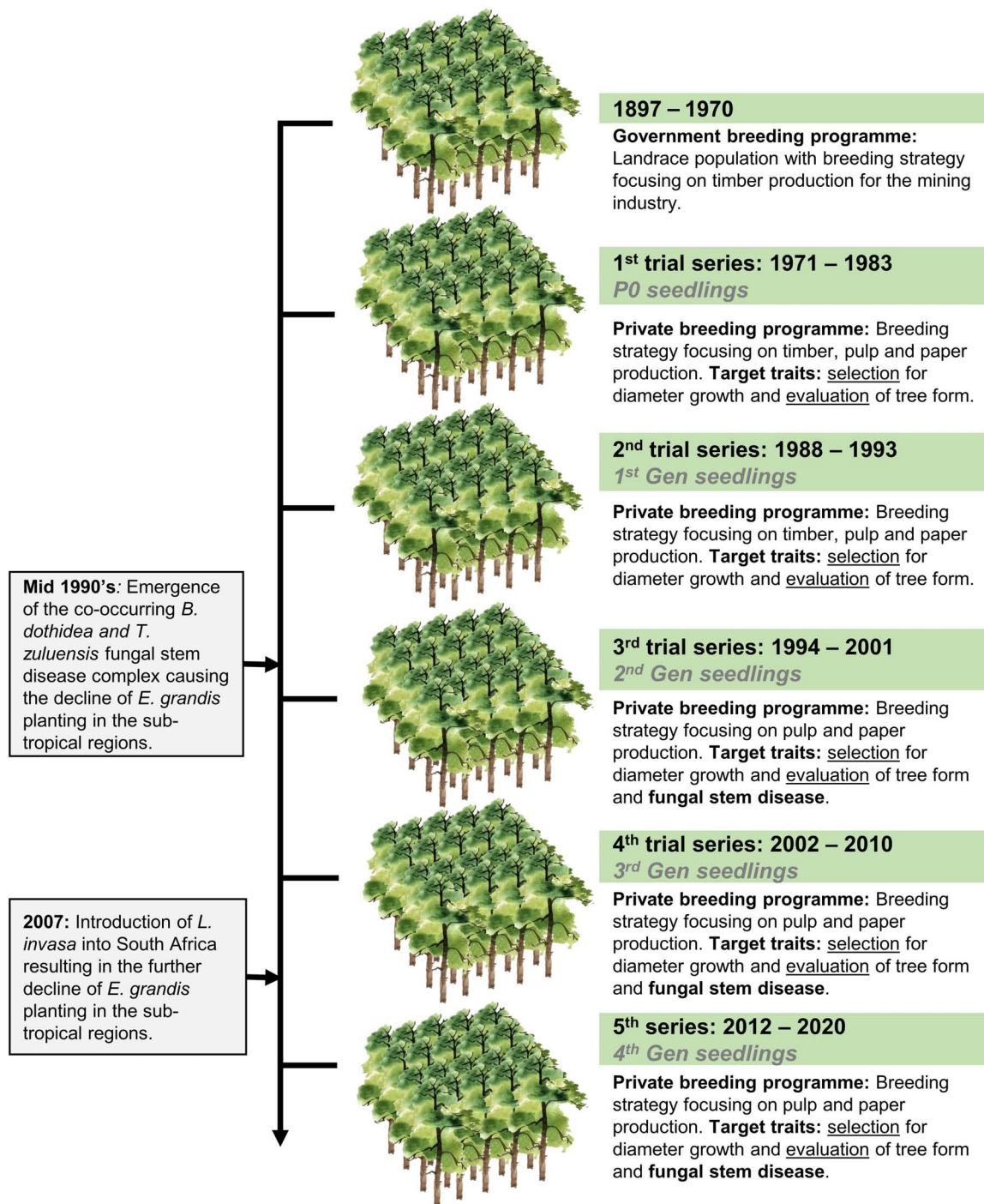


FIGURE 1 | Historical overview of *E. grandis* breeding in South Africa, including a transition from government to private breeding and introduction of major pest and pathogens. The trial series timeline, as well as the generational timeline, are shown. Selection strategies are noted for each trial series, shifting from timber to pulp and paper related traits, as well as pest and disease tolerance. Selection refers to the selection of phenotyped individuals based on their breeding values, whereas evaluation refers to the selection of individuals based on visual screening without breeding values.

where A is the relationship matrix of the random effects, σ_e^2 is the variance associated with the residuals, and σ_u^2 is the variance associated with the random effects. The assumptions of residual matrix R was relaxed to have a heterogeneous error variance

across the environments. Similarly, the assumption of the G matrix was relaxed to model full $G \times E$ and heterogeneous genetic variances at each site ($s + 1$ variance parameters), where s is the number of environments (Isik et al., 2017). Empirical breeding

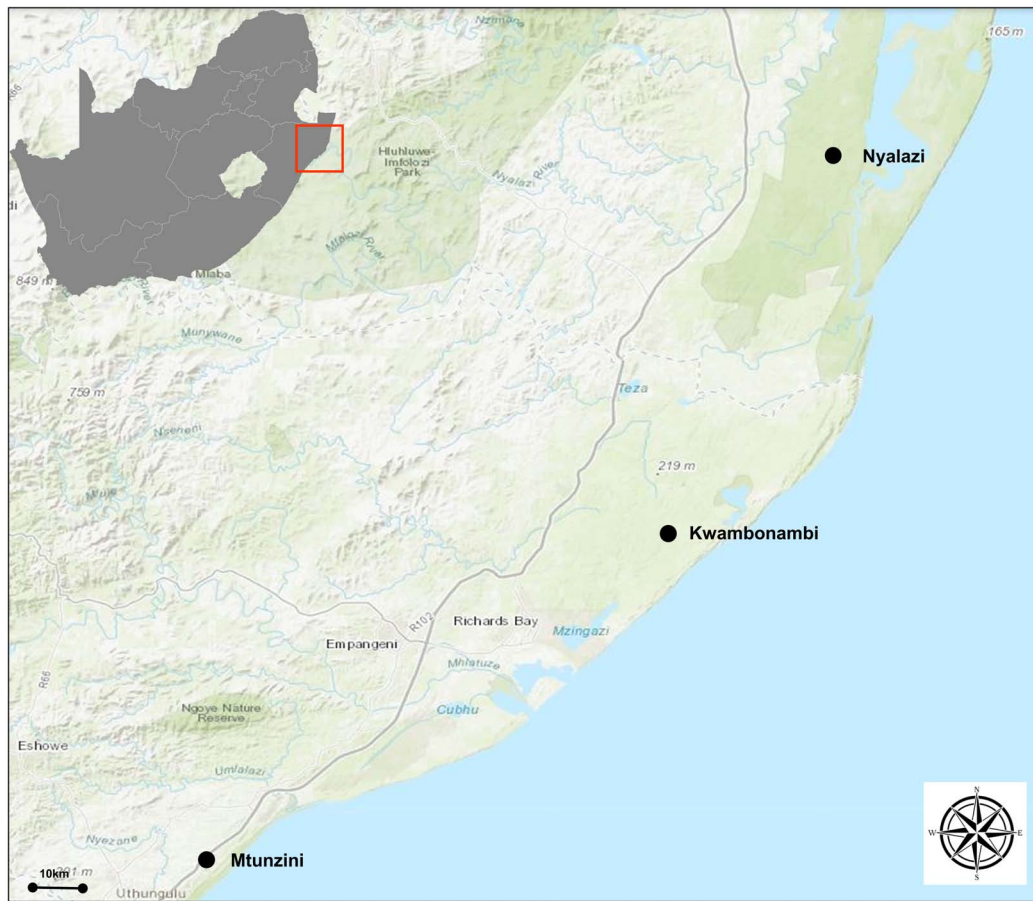


FIGURE 2 | Geographical representation of the trial sites in the KwaZulu Natal province, South Africa. The region has a sub-tropical climate. The distance (straight line) between Mtunzini and Nyalazi is 112 km. The details of the environmental conditions are in **Supplementary Table 1**. Darker shades of green indicate nature reserves.

value prediction for the half-sibs was performed by solving the mixed model equations.

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\lambda \end{bmatrix} \begin{bmatrix} \beta \\ u \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (2)$$

where A^{-1} is the inverted additive genetic relationship matrix derived from the pedigree and $\lambda = \frac{\sigma_e^2}{\sigma_u^2}$ is the shrinkage factor. The genomic relationship matrix G from the genotyped trees was computed as described in VanRaden (2008):

$$G = \frac{(Z-P)(Z-P)'}{2 \sum p_i(1-p_i)} \quad (3)$$

where Z and P are two matrices of dimension n (individuals) \times p (markers). The base pair calls were transformed into gene content values of the minor alleles at each SNP loci in each individual in matrix Z , with elements -1 (homozygote major allele), 0 (heterozygote), and 1 (homozygote minor allele). The frequencies of the genotypes were 0.584, 0.338, and 0.078, respectively. The allele frequencies in matrix P are presented as $2(p_i - 0.5)$, where p_i is the observed allele frequency at the marker i for all individuals.

The $2 \sum p_i(1-p_i)$ is the variance of alleles summed across all the loci. A ssGBLUP model was fitted using a blended relationship (H) matrix, incorporating the (G) matrix of genotyped trees that are linked to the non-genotyped trees by the half-sib pedigree (A) matrix (Legarra et al., 2009; Aguilar et al., 2010; Christensen and Lund, 2010).

The H matrix used in the ssGBLUP was formulated as follows: where u is a vector of genetic effects with variances $Var(u) = A\sigma_u^2$. Within the genetic effects (u), there are non-genotyped and (u_1) and genotyped (u_2) individuals partitioned in the A matrix as:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (4)$$

where A_{11} is the relationship matrix of non-genotyped individuals, A_{22} is the relationship matrix for the genotyped individuals, and A_{12} and its transpose A_{21} are the covariances between the genotyped non-genotyped individuals. We then replaced the u_2 genetic effects with the pedigree relationship of A_{22} with their G matrix as constructed in Eq. 3. The relationship between the non-genotyped and (u_1) and genotyped (u_2) individuals in A_{12} and A_{21} is then adjusted by the G matrix via



FIGURE 3 | Symptoms and incidence scores of *Leptocybe invasa* (*Lepto*). **(A)** Score 4 – No evidence of an attack on the leaf midrib or petiole, **(B)** Score 3 – Evidence of attack on the leaf midrib or petiole without galls (indicated by red arrows), **(C)** Score 2 – Evidence of attack on the leaf midrib or petiole with galls, and **(D)** Score 1 – Evidence of a lethal outcome of an attack on the leaf midrib or petiole with galls.

the pedigree relationship of all other individuals in the H matrix (Legarra et al., 2009):

$$H = \begin{bmatrix} A_{11} + A_{12}A_{22}^{-1}(G - A_{22})A_{22}^{-1}A_{21} & A_{12}A_{22}^{-1}G \\ GA_{22}^{-1}A_{21} & G \end{bmatrix} \quad (5)$$

The upper left corner of the H matrix is the variance of the u_1 individuals, with $\text{Var}(u_1) = [A_{11}A_{12}A_{22}^{-1}(G - A_{22})A_{22}^{-1}A_{21}] \sigma_A^2$, and $\text{Var}(u_2) = G \sigma_A^2$ and $\text{Cov}(u_1, u_2) = A_{12}A_{22}^{-1}G \sigma_A^2$. The inverse of the H matrix is:

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix} \quad (6)$$

Variance components from the ABLUP and ssGBLUP were estimated along with the heritability for diameter growth and *Lepto* and *BotryoTera* tolerance across and within the three sites.

Multivariate Analysis

A multivariate linear mixed model was fitted to estimate additive genetic correlations between three pairs of traits as described in Isik et al. (2017), following the multivariate model general design:

$$y_{n \times d} = X_{n \times (p+1)} \beta_{(p+1) \times d} + Z_{n \times r} u_{r \times d} + \epsilon_{n \times d} \quad (7)$$

where n is the number of rows of individuals and d is the number of dependent variables (traits). The design matrix X has the dimensions $n = (p1)$, where p is the number of fixed estimators, which are replication nested in location for the traits, and the additional column is added for the intercept. β is the matrix of coefficients of fixed predictor effects to be estimated with dimensions $(p1) = d$. The rows of β correspond to predictor variables, and the columns are response variables. The design matrix of Z has dimensions $n = r$, where r is the number of random effects (individual trees) per trait, and u is a $r = d$ matrix of the random effects.

The G and R variance–covariance matrices of the multivariate model were designed with the variances for the three traits on the diagonal and the covariances between the traits on the off-diagonals:

$$G = A \otimes \begin{bmatrix} \sigma_{A11}^2 & \sigma_{A12} & \sigma_{A13} \\ \sigma_{A21} & \sigma_{A22}^2 & \sigma_{A23} \\ \sigma_{A31} & \sigma_{A32} & \sigma_{A33}^2 \end{bmatrix} \quad (8)$$

$$R = I_m \otimes \begin{bmatrix} \sigma_{\epsilon11}^2 & \sigma_{\epsilon12} & \sigma_{\epsilon13} \\ \sigma_{\epsilon21} & \sigma_{\epsilon22}^2 & \sigma_{\epsilon23} \\ \sigma_{\epsilon31} & \sigma_{\epsilon32} & \sigma_{\epsilon33}^2 \end{bmatrix} \quad (9)$$

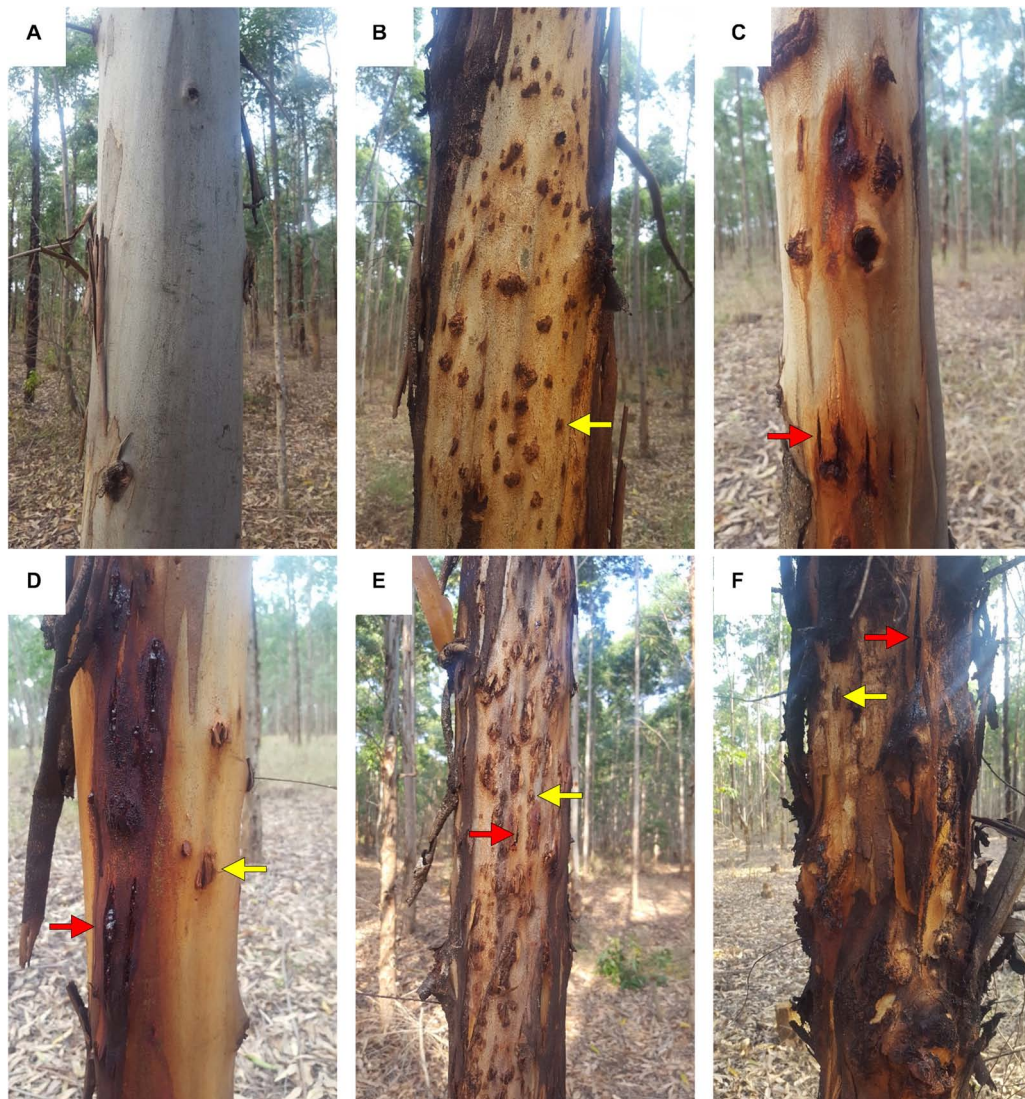


FIGURE 4 | Symptoms and incidence scores for *Botryosphaeria/Teratosphaeria* stem disease complex (*BotryoTera*). **(A)** A score of 6 represents trees with no spots/cracks or redness. **(B)** A score of 5 represents trees with *T. zuluensis* spots with redness. **(C)** A score of 4 is given for trees with *B. dothidea* cracks with redness. **(D)** A score of 3 shows a tree with *T. zuluensis* spots and *B. dothidea* cracks with redness. **(E)** A score of 2 represents trees with heavy *T. zuluensis* spots and *B. dothidea* cracks with redness. **(F)** A score of 1 represents trees with heavy *T. zuluensis* spots and *B. dothidea* cracks with redness and cankers.

where the G matrix is the direct product of the A matrix (pedigree relationship) for the ABLUP model and substituted with the H matrix for the ssGBLUP model with an unstructured, heterogeneous variance and covariance structure, where each environment has a unique genetic variance, and each pair of the environments has a unique covariance, with an $s(s-1)/2$ variance parameter (Isik et al., 2017). The R matrix is the direct product of the identity matrix (I_m) with m dimensions, m is the number of genotypes with variance $\sigma_{\epsilon_1}^2$ for diameter growth, $\sigma_{\epsilon_2}^2$ for *BotryoTera*, and $\sigma_{\epsilon_3}^2$ for *Lepto* and their covariances nested within.

The construction of the expected additive (A matrix) and the realized genomic (G) was calculated using the package *synbreed* 0.10-2 (Wimmer et al., 2012) in the R environment v3.5.3. The blended genetic relationships and its inverse were

obtained using scripts according to Isik et al. (2017). All the statistical models were performed using ASReml software v4.1 (Gilmour et al., 2015).

Expected Direct and Indirect Genetic Gains

The direct genetic gains for *diameter* growth and *Lepto* and *BotryoTera* tolerance were calculated from the ABLUP and ssGBLUP models breeding value predictions. The selection differential was based on the top 10% of individuals for direct selection. The indirect responses of the remaining traits were calculated based on the ranking of the direct selections. The percentage expected genetic gains were calculated as the fraction of the selection differential over the population mean.

RESULTS

Genetic Parameters

To assess the increased accuracy of the ssGBLUP model, we compared the heritability estimates from ssGBLUP with those from ABLUP analysis. The ssGBLUP model generally produced lower heritability estimates compared to the ABLUP model for the three sites (Table 1). The exception was the heritability estimates for *BotryoTera* tolerance in Kwambonambi and Nyalazi, which were higher for ssGBLUP (0.45 vs. 0.29 and 0.11 vs. 0.08, respectively). Overall, the Kwambonambi site produced the highest heritability values ranging from 0.29 to 0.63 (ABLUP) and from 0.45 to 0.70 (ssGBLUP) across the traits (Table 1). In contrast, the heritability estimates for *Lepto* tolerance from the ABLUP and ssGBLUP models were the highest at 0.71 and second highest at 0.38, respectively, in Nyalazi, while the estimates for diameter growth and *BotryoTera* tolerance at the Nyalazi site were reasonably low, ranging from 0.07 to 0.11 for the ABLUP and ssGBLUP models, respectively (Table 1). The overall heritability estimates across sites were higher for the ABLUP model with *Lepto* tolerance moderately high at 0.54, diameter

growth at 0.33, and *BotryoTera* tolerance at 0.23 (Table 2). The heritability estimates with the ssGBLUP across sites were lower with *Lepto* tolerance at 0.36, diameter growth at 0.25, and *BotryoTera* tolerance at 0.23 (Table 2). The heritability estimates for ssGBLUP may be more accurate due to the blended pedigree relationship matrix increased precision.

ssGBLUP Additive and Type-B Genetic Correlations

The additive genetic correlations of diameter growth and *Lepto* tolerance estimated with the ssGBLUP model was high at 0.78 (Table 3, Eq. 7). In contrast, the additive genetic correlation of diameter growth and *BotryoTera* tolerance was moderate at -0.38 . The additive genetic correlation for *BotryoTera* and *Lepto* tolerance was also moderate at -0.47 (Table 3). These results suggest that tandem improvement of diameter growth and *Lepto* tolerance is possible, but they predict a negative response in *BotryoTera* tolerance, which presents a challenge to breeders. The overall Type-B genetic correlation (Eq. 7) was high, ranging from 0.77 to 0.81 for the three traits associated with small standard errors (Table 4), suggesting low $G \times E$ interactions across the sites.

Trait Performance Across Site and Generations

Diameter growth and the *Lepto* incidence scores resembled a normal distribution (Supplementary Figure 1). *BotryoTera* incidence scores had a high frequency of score 6, representing

TABLE 1 | Site-specific variance components and genetic parameters estimated using the ABLUP and ssGBLUP mixed models for diameter growth, *BotryoTera* and *Lepto* tolerance.

| | σ_u^2 (se) | σ_e^2 (se) | h^2 (se) |
|--------------------------|-------------------|-------------------|---------------|
| Diameter | | | |
| ABLUP | | | |
| Mtunzini | 6.655 (0.281) | 2.360 (0.655) | 0.36 (0.092) |
| Kwambonambi | 13.193 (0.662) | 8.250 (1.945) | 0.64 (0.129) |
| Nyalazi | 11.928 (0.547) | 0.884 (0.579) | 0.07 (0.048) |
| ssGBLUP | | | |
| Mtunzini | 6.670 (0.277) | 1.620 (0.504) | 0.24 (0.072) |
| Kwambonambi | 13.592 (0.682) | 7.852 (1.487) | 0.58 (0.092) |
| Nyalazi | 11.958 (0.552) | 0.779 (0.582) | 0.07 (0.048) |
| <i>BotryoTera</i> | | | |
| ABLUP | | | |
| Mtunzini | 1.450 (0.055) | 0.424 (0.115) | 0.29 (0.0752) |
| Kwambonambi | 2.334 (0.099) | 0.115 (0.203) | 0.30 (0.0823) |
| Nyalazi | 1.411 (0.059) | 0.109 (0.056) | 0.08 (0.0393) |
| ssGBLUP | | | |
| Mtunzini | 1.447 (0.053) | 0.222 (0.077) | 0.15 (0.052) |
| Kwambonambi | 2.404 (0.110) | 1.088 (0.227) | 0.45 (0.083) |
| Nyalazi | 1.418 (0.060) | 0.154 (0.073) | 0.11 (0.051) |
| <i>Lepto</i> | | | |
| ABLUP | | | |
| Mtunzini | 0.454 (0.017) | 0.161 (0.039) | 0.36 (0.080) |
| Kwambonambi | 0.762 (0.035) | 0.524 (0.105) | 0.70 (0.118) |
| Nyalazi | 0.764 (0.037) | 0.542 (0.112) | 0.71 (0.125) |
| ssGBLUP | | | |
| Mtunzini | 0.452 (0.016) | 0.110 (0.026) | 0.24 (0.055) |
| Kwambonambi | 0.770 (0.033) | 0.538 (0.070) | 0.70 (0.072) |
| Nyalazi | 0.744 (0.031) | 0.281 (0.049) | 0.38 (0.059) |

The residual variance (σ_e^2), additive genetic variance (σ_u^2), narrow-sense heritability (h^2), and their standard errors (se) are shown.

TABLE 2 | Overall variance components and genetic parameters across the three sites for solving ABLUP and ssGBLUP mixed models for diameter growth, *BotryoTera*, and *Lepto* tolerance.

| | σ_u^2 (se) | σ_e^2 (se) | h^2 (se) |
|-------------------|-------------------|-------------------|--------------|
| ABLUP | | | |
| Diameter | 10.581 (0.314) | 3.450 (0.720) | 0.33 (0.063) |
| <i>BotryoTera</i> | 1.732 (0.044) | 0.407 (0.092) | 0.24 (0.051) |
| <i>Lepto</i> | 0.659 (0.021) | 0.357 (0.059) | 0.54 (0.077) |
| ssGBLUP | | | |
| Diameter | 10.729 (0.313) | 2.733 (0.469) | 0.26 (0.040) |
| <i>BotryoTera</i> | 1.755 (0.046) | 0.396 (0.071) | 0.23 (0.038) |
| <i>Lepto</i> | 0.655 (0.017) | 0.238 (0.024) | 0.36 (0.032) |

The residual variance (σ_e^2), additive genetic variance (σ_u^2), narrow-sense heritability (h^2), and its standard error (se) are presented.

TABLE 3 | Additive genetic correlations (r_g) of diameter growth, *BotryoTera*, and *Lepto* tolerance based on ABLUP and ssGBLUP models with standard errors in the parenthesis.

| | <i>BotryoTera</i> | <i>Lepto</i> |
|-------------------|-------------------|-----------------|
| ABLUP | | |
| Diameter | -0.46 (0.116) | 0.81 (0.054) |
| <i>BotryoTera</i> | | -0.47 (0.111) |
| ssGBLUP | | |
| Diameter | -0.38 (0.106) | 0.78 (0.055) |
| <i>BotryoTera</i> | | -0.47 (0.089) |

TABLE 4 | Overall Type-B genetic correlation (r_B) across sites for diameter growth, *BotryoTera*, and *Lepto* tolerance based on ABLUP and ssGBLUP models with standard errors in the parenthesis.

| | r_B (se) |
|-------------------|--------------|
| ABLUP | |
| Diameter | 0.90 (0.096) |
| <i>BotryoTera</i> | 0.99 (0.000) |
| <i>Lepto</i> | 0.88 (0.055) |
| ssGBLUP | |
| Diameter | 0.80 (0.140) |
| <i>BotryoTera</i> | 0.81 (0.147) |
| <i>Lepto</i> | 0.77 (0.072) |

uninfected stems, and Kwambonambi has a high frequency of score 3 (**Supplementary Figure 1**). The latter may be ascribed to the second-generation families' higher susceptibility (**Figure 5B** and **Supplementary Figure 2**). The Kwambonambi site had the lowest mean *BotryoTera* tolerance compared to the Nyalazi and Mtunzini (**Figure 5E**). The average diameter growth improved by 3.2% from the third to the fourth generation (**Figure 5A**), whereas *Lepto* tolerance improved by 3.6% (**Figure 5C**). The improvement in diameter growth is driven by recurrent selection over the generations with *Lepto* tolerance benefiting from

its strong additive genetic correlation with diameter growth (**Table 3**). There was a 13.3% improvement of *BotryoTera* tolerance from the second to the third generation; however, it was unchanged from the third to the fourth generation (**Figure 5B**). The apparent absence in genetic gain for *BotryoTera* tolerance from the third to the fourth generation is in part due to the moderately negative genetic correlation with diameter growth (**Table 3**). The above results suggest that a revised breeding strategy is needed to improve the three traits simultaneously.

Correlated Response Based on ssGBLUP Breeding Values

The direct genetic gains estimated for diameter growth and *Lepto* tolerance were 12.4% and 24.7%, respectively, with *BotryoTera* at 9.8% (**Table 5**). There is an indirect loss of 3.4% in *BotryoTera* tolerance and a gain of 10.0% in *Lepto* tolerance when selecting for diameter growth. Direct selection for *BotryoTera* tolerance would result in an expected indirect loss of 5.6% for diameter growth and 6.5% for *Lepto* tolerance. However, direct selection of *Lepto* tolerance would result in an expected gain of 6.0% for diameter growth and loss of 3.8% in *BotryoTera* tolerance (**Table 5**). Together, these results illustrate the challenge of achieving genetic gains for all three of these traits and the need for customized breeding strategies to deal with this challenge.

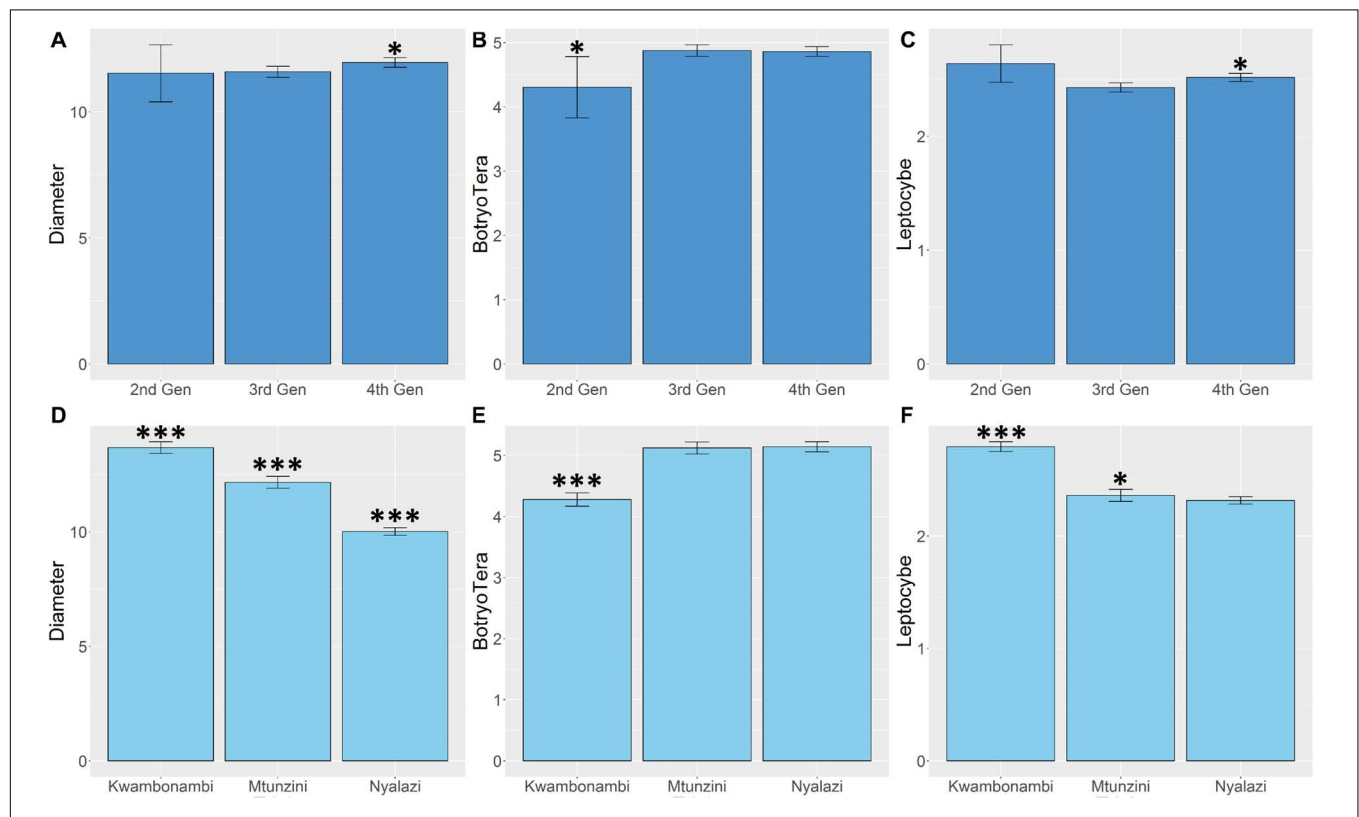


FIGURE 5 | Marginal trait means with error bars indicating the 95% confidence interval. (A) Mean diameter growth (cm) for families in the three sites. (B) The mean *Lepto* tolerance score for families in the three sites. (C) The mean *BotryoTera* tolerance score for families in the three sites. (D) Mean diameter growth (cm) for families in the three generations. (E) The mean *Lepto* tolerance score for families in the three generations. (F) The mean *BotryoTera* tolerance score for families in the three generations. Student *t*-test was performed to assess the significant difference between the means, $p < 0.05$ (*) and $p < 0.001$ (***).

TABLE 5 | Expected genetic gains (%) based on the top 10% selected individuals in the population for diameter growth, *Botryotera*, *Lepto* tolerance, and the indirect response in the expected genetic gains of the paired traits.

| | | Diameter | <i>Botryotera</i> | <i>Lepto</i> |
|-------------------|-------------------|-------------|-------------------|--------------|
| ABLUP | | | | |
| Direct response | | 15.1 | 8.9 | 32.2 |
| Indirect response | Diameter | 15.1 | −3.5 | 9.9 |
| | <i>Botryotera</i> | −6.1 | 9.3 | −8.0 |
| | <i>Lepto</i> | 5.9 | −1.9 | 26.4 |
| ssGBLUP | | | | |
| Direct response | | 12.4 | 9.8 | 24.7 |
| Indirect response | Diameter | 12.4 | −3.4 | 10.0 |
| | <i>Botryotera</i> | −5.6 | 10.1 | −6.5 |
| | <i>Lepto</i> | 6.0 | −3.8 | 20.9 |

The bold diagonals are the direct response with the off-diagonal as the indirect responses.

DISCUSSION

Pest and pathogens are significant risk factors in forest plantations (Wingfield et al., 2015). These risk factors are highlighted in African agroforestry systems affecting indigenous and natural forests (Graziosi et al., 2020). Mitigation of these risk factors will require recognizing the parallels and synergies in management methods between pest and pathogen studies (Jactel et al., 2020) and integrating system genetic and systems biology (Naidoo et al., 2019) particularly in this genomic era (Naidoo et al., 2014). The continued improvement of economic traits such as volume growth, density, and pulp yield in the context of pest and pathogen challenges is vital. Here, we combined phenotypic data for a large half-sib breeding trial with genotypic data for a subset of siblings in a single-step genomic BLUP approach to estimate genetic parameters and response to selection for diameter growth and *Botryotera* and *Lepto* tolerance in *E. grandis* breeding population. We also proposed a practical genomic selection breeding strategy that is likely to improve all three traits in *E. grandis*. One of the study strengths was the availability of replicated trials with *Botryotera* infections and *Lepto* infestation across all three sites.

Furthermore, the study benefited from planting pedigree-linked families from three successive generations in the same space and time. Therefore, the trials provided an opportunity to evaluate the outcomes of three different artificial selection regimes applied in successive generations. A limitation was the inability to score *B. dothidea* and *T. zuluensis* infections separately, which we mitigated by developing a combined phenotypic score (Figure 3). Diameter growth and *Botryotera* and *Lepto* tolerance had moderate heritability estimates (0.25–0.36, Table 2). Diameter growth and *Lepto* tolerance had a strong positive additive genetic correlation. However, both were negatively correlated with *Botryotera* tolerance, though the correlations were not strong. This presents a challenge to achieve genetic gains in all three traits simultaneously.

Genetic Parameters for Diameter Growth and *Lepto* and *Botryotera* Tolerance

Coefficients of relationship from pedigree data are expectations and do not represent the actual genome shared between relatives, estimated from various allelic frequency parameters (Forni et al., 2011). Forest trees with deep full-sib pedigrees have estimated coefficients of relationship that are much closer to the actual genetic relationships (Batholome et al., 2016; Chen et al., 2018). However, more precise coefficients of relationship are estimated using DNA markers such as SNPs (Habier et al., 2007; Hayes et al., 2009). When expected genetic relationships are combined with the genome estimated relationships, this precision can be extrapolated to the A matrix with the blended H matrix used in ssGBLUP analyses (Legarra et al., 2009; Aguilar et al., 2010). Half-sib pedigree relationships do not include cryptic genetic relationships in the population, in some instances leading to biased estimation of additive genetic variances (Ratcliffe et al., 2017).

In this study, we generally observed lower heritability estimates from ssGBLUP compared to ABLUP (Table 2). Lower additive genetic correlation estimates were also observed for ssGBLUP compared to ABLUP (Table 3). Luo et al. (2014) presented heritability estimates of *Lepto* tolerance in *E. camaldulensis* and *E. tereticornis* breeding populations in China of 0.54 and 0.52, respectively. da Silva et al. (2020), also presented heritability estimated from multiple *Eucalyptus* species ranging from 0.27 to 0.68, with *E. grandis* at 0.58. These heritability estimates are similar to what we obtained in our study at 0.54 for *E. grandis* (Table 2). The *Lepto* tolerance scores in the study by Luo et al. (2014) were based on the proportion of the canopy affected, with a score of 0 indicating no symptoms on the canopy and a score of 4 meaning greater than 75% of the canopy affected (Thu et al., 2009).

In contrast, our scoring system was not based on canopy proportions, but rather the severity of gall formation with a score of 4 indicating no evidence of gall formation and a score of 1 indicating lethal outcome from gall formation in both mid-ribs and petioles of the leaves (Figure 3). Luo et al. (2014) reported a moderately negative genetic correlation between tree height (at 9 months) and *Lepto* susceptibility in *E. camaldulensis* at −0.33 and for *E. tereticornis* at −0.47. Due to the inverted scores used in our study, we report a positive genetic correlation (0.78) with diameter growth at 48 months (Table 3). These results suggest that vigorous tree growth is positively related to tolerance to *L. invasa*. Plant growth regulators are well-characterized phytohormones involved in influencing plant development and abiotic stress responses (Wani et al., 2016) and pest tolerance (Harun-or-Rashid and Chung, 2017). There is evidence to suggest that the microbiome of the maternal environment may affect the performance of their progeny and tolerance to pathogens in *E. grandis* (Vivas et al., 2017). A study to characterize the relationship of maternal and/or progeny microbiomes, phytohormones, and their interactions, on superior tree growth and health, is warranted.

Genotype-by-Environment Interaction and Trait Performance

The mean annual precipitation of the three sites in the subtropical region of South Africa decreases from South to North, tracking the increase in the mean annual temperature maximum (Figure 2). Therefore, Nyalazi in the North is on average warmer and drier compared to Mtunzini in the South, which is on average colder and wetter, whereas Kwambonambi has mid-range environmental conditions (Supplementary Table 1). The pairwise Type-B genetic correlation for diameter growth and *Lepto* and *Botryotera* tolerance across the sites ranged from 0.77 to 0.81 (Table 4), indicating low $G \times E$ interaction. The Nyalazi trial was surrounded by a commercial stand of *E. grandis* \times *E. camaldulensis* ($G \times C$) clone that was highly susceptible to *L. invasa*. The $G \times C$ hybrid genotype has been shown in the literature to be susceptible to *L. invasa* (Thu et al., 2009; Luo et al., 2014). The $G \times C$ clone planted in the Nyalazi site had an increased infestation of *L. invasa* translating into the high frequency of *Lepto* tolerance score 2 in the trial and much lower frequency of *Lepto* tolerance score 3 and 4 (Supplementary Figure 1). In Mtunzini, there was also an increased frequency of *Lepto* score 2; however, the trial was surrounded by a tolerant *E. grandis* \times *E. urophylla* ($G \times U$) clone (Supplementary Figure 1). There are above-average actively growing shoots in Mtunzini due to its favorable environmental conditions (Supplementary Table 1). These actively growing shoots are targets for *L. invasa* infestation. The heritability estimates of *Lepto* tolerance in Mtunzini and Nyalazi were adjusted lower from 0.35 to 0.24 and 0.71 to 0.38, respectively, by the ssGBLUP model (Table 1). It is not clear why the heritability correction in Nyalazi was so significant compared to that in Mtunzini.

In Kwambonambi, the mid-range environmental conditions to Mtunzini and Nyalazi, which was also surrounded by a tolerant $G \times U$ clone, *Lepto* tolerance showed similar heritability estimates between ABLUP (0.69) and ssGBLUP (0.70) and for diameter growth ABLUP (0.63) and ssGBLUP (0.58) (Table 1). The similar heritability estimates in Kwambonambi of diameter growth and *Lepto* tolerance may result from their relatively high positive additive genetic correlation. The estimated marginal means for diameter growth and *Lepto* tolerance in Kwambonambi further support this relationship (Figures 5D,F).

There is an increased incidence of *Botryotera* tolerance score 3 in Kwambonambi (Supplementary Figure 1), resulting from the increased susceptibility from the second-generation families (Supplementary Figure 2). *Botryotera* appeared as a fungal stem disease in the mid- to late 1990s, which means that the first-generation parents (second-generation families) were selected in the absence of the *Botryotera* disease explaining the higher susceptibility of the second generation families. The environmental conditions at the Kwambonambi site are optimal for diameter growth, and, due to the negative correlation with *Botryotera* tolerance, there was high susceptibility to *Botryotera* in Kwambonambi (Figure 5E). Diameter growth and *Lepto* and *Botryotera* tolerance in the Kwambonambi site, which is the mid-range of Nyalazi and Mtunzini environmental conditions, seem

to reflect the trait performances, corresponding to their additive genetic correlation.

Generational Performance for Diameter Growth and *Lepto* and *Botryotera* Tolerance

Recurrent selection in tree breeding ensures the gradual improvement of target economic traits over generations. Such efforts are under threat from pest and pathogen pressures as well as climate change (Wingfield et al., 2015). Reversing the decline of *E. grandis* in the subtropical region of South Africa due to *L. invasa* gall wasp and the co-occurrence of *B. dothidea* and *T. zuluensis* fungal stem disease is vital. *Botryotera* fungal stem disease was discovered and described in South Africa in the early to mid-1990s (Smith et al., 1994; Wingfield et al., 1996). This meant that selections or evaluations in the government landrace breeding populations did not involve *Botryotera* tolerance until the first generation in the 2nd trial series and onwards in the private breeding population (Figure 1), evidenced by the high *Botryotera* incidence score 3 (Supplementary Figure 2) of the second-generation families in particular in the Kwambonambi site (Supplementary Figure 1). Evaluation for *Botryotera* tolerance in the second generation resulted in the increased tolerance in the third generation and maintained in the fourth generation (Figure 5B). When looking at the high frequency of *Botryotera* score 6 in Supplementary Figures 1, 2, it does suggest that the evaluation strategy has had a limited role to play in improving *Botryotera* tolerance, because this trait seems to have plateaued in the last generations. The limitation of the evaluation strategy for *Botryotera* tolerance is that selection was only performed within families already selected for diameter growth and further compounded by the fact that *Botryotera* tolerance is negatively correlated with diameter growth.

Leptocybe invasa was reported in South Africa in 2007 (Neser et al., 2007), coinciding with the third generation tested in the 4th trial series (Figure 1). *Leptocybe* appeared when the trial series was at age 5. The canopies were already inaccessible for scoring and selecting *Lepto* tolerance for the fourth generation (Figure 1). The indirect improvement of *Lepto* tolerance from the third to the fourth generation is due to the strong positive additive genetic correlation with diameter growth (Figure 5C). This study showed that the recurrent selection strategy successfully improved diameter growth and indirectly improved *Lepto* tolerance, with limited impact on *Botryotera* tolerance.

Proposed Selection Strategies for Diameter Growth and *Lepto* and *Botryotera* Tolerance

Eucalypts, including *E. grandis*, are currently experiencing a decline, mainly due to pest and pathogen pressures for commercial deployment and breeding populations such as *Puccinia psidii* (Silva et al., 2013), *L. invasa* (da Silva et al., 2020), *T. zuluensis* (Wingfield et al., 1996; Aylward et al., 2019), and *B. dothidea* (Smith et al., 1996; Marsberg et al., 2017). This study offers opportunities to revise historical evaluation and selection

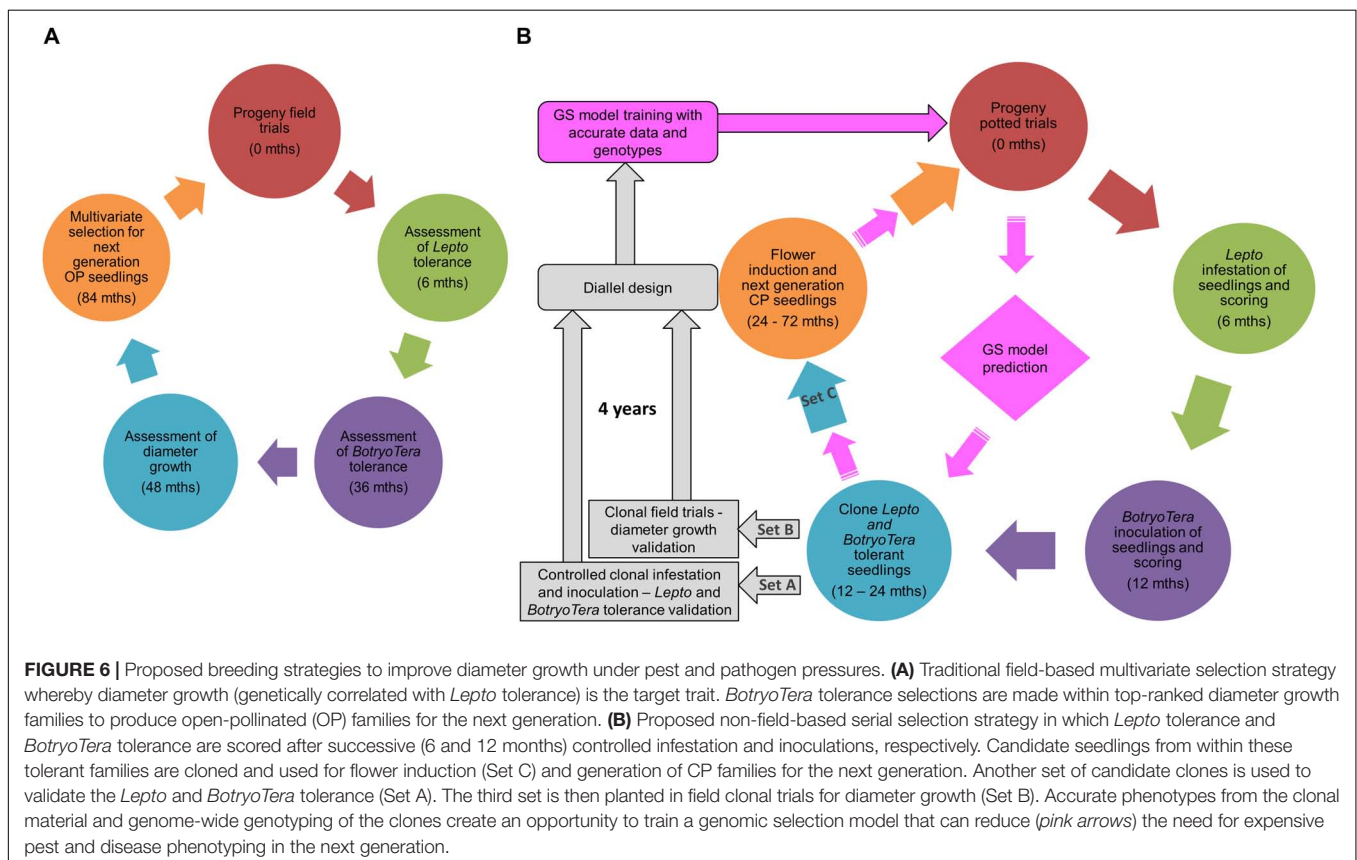
strategies to improve diameter growth and *Botryotera* and *Lepto* tolerance. Testing all these pedigree-linked *E. grandis* generations in the same space and time has highlighted the successes and challenges of traditional evaluation and selection strategies and their direct and indirect impact on economic traits over the generations as new pests and pathogens emerge. First, pests and pathogens may appear during a growth stage within a breeding cycle when trees cannot be effectively scored and selected. Second, pests and pathogens affect different parts of the tree, young leaves (early in the growth cycles), and stem (later in the growth cycles); therefore, the correct timing of scoring is crucial. Third, although present, pests and pathogens may differ in their infestation and infection severity due to many factors, leading to highly varying levels of challenge and incomplete expression of tolerance or susceptibility. Fourth, the emergence of pests and pathogens sometimes may reveal inadequacies of already established selection strategies, thereby requiring revision, as is the case for *Botryotera*.

A multivariate approach to deal with these challenges requires an understanding of the traits additive genetic correlations. Such a strategy would require turning over a generation in which all three traits were measured on each tree to estimate their between- and within-family breeding values. The challenge with field trials is that there are often difficulties to score pest and pathogen tolerance accurately, as discussed. Breeders may adopt a multivariate approach to primarily select for diameter growth and indirectly for *Lepto* tolerance and then only consider

selecting *Botryotera* tolerant individuals from high ranked families (Figure 6A).

Circumventing field trials and the inconsistency of pest infestations or pathogen infections, tree breeders may consider a proposed serial selection strategy with genomic selection and controlled pollination in potted trials (Figure 6B). This approach would require the integration of nursery and field phenotypes to develop a more accurate GS model. Such an approach was demonstrated in *Populus deltoids* for tree height to accelerating its breeding strategies (Alves et al., 2020). The proposed GS approach in this involves challenging potted families with *L. invasa* and scoring *Lepto* tolerance 6 months after potting and then advancing the most tolerant individuals across families for *Botryotera* tolerance scoring at 12 months after potting. The best individuals from the top *Lepto* and *Botryotera* tolerant families are then cloned to validate the pest and pathogen tolerance (Set A).

Meanwhile, the second set of ramets from the same clones (Set B) is planted in field trials to validate the expected correlated diameter growth response, while the third set of ramets (Set C) are subjected to flower induction to produce control-pollinated next-generation families. The clonal phenotypic data can be used together with genome-wide genotyping to train a genomic selection model for implementation (pink arrows in Figure 6B). Genomic estimated breeding values and genomic relationship matrices will inform the control pollination (diallel in the potted orchard) (Munoz et al., 2014; Li et al., 2019). This approach



should increase the selection intensity and reduce the need for costly controlled pest and pathogen challenges, thereby fast-tracking clonal tests and producing next-generation control-pollinated (CP) seedlings (with breeding value predictions for all three traits) to improve gains per unit time over what can be achieved in a traditional open-pollinated (OP) field testing approach.

CONCLUSION

Diameter growth and pest and pathogen tolerance are essential components of sustainable plantation forestry. Therefore, a multivariate selection approach informed by their additive genetic correlations is key to improving genetic gains in these traits simultaneously. This study shows that evaluation and selection strategies implemented for *E. grandis* over the past three generations have succeeded in improving diameter growth and indirectly *Lepto* tolerance, while limited gain was achieved for *BotryoTera* tolerance. We proposed an alternative to the traditional field-based multivariate strategy, which has many challenges mainly limited by the reliability of assessing pest infestations and pathogen infections in the field. The proposed serial genomic selection strategy involves controlled infestations with *Lepto* and inoculations with *BotryoTera* of cloned families in pots to achieve validated and accurate tolerance scores and diameter growth measurements from clonal field trials. This approach will ensure a reliable multivariate genomic selection training and development to exploit the additive genetic correlations void phenotyping challenges with field trials. The proposed genomic selection strategy, possibly via ssGBLUP (Miszta et al., 2013), would be a feasible approach to improve diameter growth and *Lepto* and *BotryoTera* tolerance in *E. grandis*.

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

MM carried out the experimental design, data collection, data analysis, and drafting of this manuscript as part of his Ph.D. thesis. FI assisted with the modeling and scripts of the data analysis tools. FI, GH, and AM helped with the data interpretation and supervision. All authors have read and approved the final manuscript.

FUNDING

The authors acknowledge financial support from the National Research Foundation (NRF) Bioinformatics and Functional Genomics Program (BFG) (grant UID 86936 and 97911), the Technology and Human Resources for Industry Program (THRIP) (grants 80118 and 964134), and Mondi South Africa.

ACKNOWLEDGMENTS

We thank Mondi South Africa Forests for the use of their *E. grandis* breeding population as well the use of facilities and technical support to generate phenotypic data. Mondi teams did the diameter at breast height measurements and the pest and pathogen scoring. SNP genotyping was supported through the Forest Molecular Genetics (FMG) Program at the University of Pretoria. The authors acknowledge North Carolina State University for training on R and ASReml analysis tools.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.638969/full#supplementary-material>

REFERENCES

- Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., and Lawlor, T. J. (2010). Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 92, 743–752. doi: 10.3168/jds.2009-2730
- Alves, F. C., Balmant, K. M., Resende, M. F. R. Jr., Kirst, M., and De Los Campos. (2020). Accelerating forest tree breeding by integrating genomic selection and greenhouse phenotyping. *Plant Genome* 13:e20048.
- Aylward, J., Roets, F., Dreyer, L. L., and Wingfield, M. J. (2019). Teratosphaeria stem canker of *Eucalyptus*: two pathogens, one devastating disease. *Mol. Plant Pathol.* 20, 9–19.
- Batholome, J., Van Heerwaarden, J., Isik, F., Boury, C., Vidal, M., Plomion, C., et al. (2016). Performance of genomic prediction within and across generations in maritime pine. *BCM Genomics* 17:604. doi: 10.1186/s12864-12016-12879-12868
- Bihon, W., Slippers, B., Burgess, T. I., Wingfield, M. J., and Wingfield, B. D. (2011). Sources of *Diplodia pinea* endophytic infections in *Pinus patula* and *P. radiata* seedlings in South Africa. *Forest Pathol.* 41, 370–375. doi: 10.1111/j.1439-0329.2010.00691.x
- Burdon, R. D. (1977). Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. *Silvae Genet.* 26, 168–175.
- Burgess, T. I., Tan, Y. P., Garnas, J., Edwards, J., Scarlett, K. A., Shuttleworth, L. A., et al. (2019). Current status of the Botryosphaeriaceae in Australia. *Australasian Plant Pathol.* 48, 35–44. doi: 10.1007/s13313-018-0577-5
- Bush, S. J., Dittrich-Schroder, G., Naser, S., Gevers, C., Baffoe, K. O., Slippers, B., et al. (2018). First record of *Quadrastichus mendeli*, a parasitoid of *Leptocybe invasa*, in South Africa. *Southern Forests J. Forest Sci.* 80, 275–277. doi: 10.2989/20702620.2017.1318347
- Cappa, E. P., De Lima, B. M., Da Silva-Junior, O. B., Garcia, C. C., Mansfield, S. D., and Grattapaglia, D. (2019). Improving genomic prediction of growth and wood traits in *Eucalyptus* using phenotypes from non-genotyped trees by single-step GBLUP. *Plant Sci.* 284, 9–15. doi: 10.1016/j.plantsci.2019.03.017
- Cetinkol, O. P., Smith-Moritz, A. M., Cheng, G., Lao, J., George, A., Hong, K., et al. (2012). Structural and chemical characterization of hardwood from tree

- species with applications as bioenergy feedstocks. *PLoS One* 7:e52820. doi: 10.1371/journal.pone.0052820
- Chang, R. L., Arnold, R. J., and Xhou, X. D. (2012). Association between enzyme activity levels in *Eucalyptus* clones and their susceptibility to the gall wasp, *Leptocybe invasa*, in South China. *J. Trop. Forest Sci.* 24, 256–264.
- Chen, S. F., Pavlic, D., Roux, J., Slippers, B., Xie, Y. L., Wingfield, M. J., et al. (2011). Characterization of *Botryosphaeriaceae* from plantation-grown *Eucalyptus* species in South China. *Plant Pathol.* 60, 739–751. doi: 10.1111/j.1365-3059.2011.02431.x
- Chen, Y., and Lubberstedt, T. (2010). Molecular basis of trait correlation. *Trends Plant Sci.* 15, 454–461. doi: 10.1016/j.tplants.2010.05.004
- Chen, Z.-Q., Baisan, J., Pan, J., Karlsson, B., Anderson, B., Westin, J., et al. (2018). Accuracy of genomic selection for growth and wood quality traits in two control pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BCM Genomics* 19:946. doi: 10.1186/s12864-018-15256-y
- Christensen, O. F., and Lund, M. S. (2010). Genomic prediction when some animals are not genotyped. *Genet. Selection Evol.* 42:2. doi: 10.1186/1297-9686-1142-1182
- Cortinas, M. N., Koch, N., Thain, J., Wingfield, B. D., and Wingfield, M. J. (2004). First record of the *Eucalyptus* stem canker pathogen, *Coniothyrium zuluense* from Hawaii. *Aust. Plant Pathol.* 33, 309–312. doi: 10.1071/ap04015
- Crous, P. W., Groenewald, J. Z., Summerell, B. A., Wingfield, B. D., and Wingfield, M. J. (2009). Co-occurring species of *Teratosphaeria* on *Eucalyptus*. *Persoonia Mol. Phylogeny Evol. Fungi* 22, 38–48. doi: 10.3767/003158509x242333
- da Silva, P. H. M., Junqueira, L. R., De Araujo, M. J., Wilcken, C. F., Moraes, M. L. T., and De Paula, R. C. (2020). Susceptibility of eucalypt taxa to a natural infestation by *Leptocybe invasa*. *New Forests* 51, 753–763. doi: 10.1007/s11056-019-09758-1
- Devappa, R. K., Rakshit, S. K., and Dekker, R. F. H. (2015). Forest biorefinery: Potential of poplar phytochemicals as value-added co-products. *Biotechnol. Adv.* 33, 681–716. doi: 10.1016/j.biotechadv.2015.02.012
- Dittrich-Schroder, G., Harney, M., Neser, S., Joffe, T., Bush, S., Hurley, B. P., et al. (2014). Biology and host preference of *Selitrachodes nesi*: a potential biological control agent of the *Eucalyptus* gall wasp, *Leptocybe invasa*. *Biol. Control* 78, 22–41.
- Dittrich-Schroder, G., Wingfield, M. J., Hurley, B. P., and Slippers, B. (2012). Diversity in *Eucalyptus* susceptibility to the gall-forming wasp *Leptocybe invasa*. *Agric. Forest Entomol.* 4, 419–427.
- Durand, N., Rodrigues, J. C., Mateus, E., Boavida, C., and Branco, M. (2011). Susceptibility variation in *Eucalyptus* spp. in relation to *Leptocybe invasa* and *Ophelimus maskelli* (Hymenoptera: Eulophidae), two invasive gall wasps occurring in Portugal. *Silva Lusitana* 19, 19–31.
- Epstein, L., Sukhwinder, K., and Vanderghenst, J. S. (2008). Botryosphaeria-related dieback and control investigated in noncoastal California grapevines. *Calif. Agric.* 62, 161–166. doi: 10.3733/ca.v062n04p161
- Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*. London: Longman Group.
- Forni, S., Aguilar, I., and Misztal, I. (2011). Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genet. Selection Evol.* 43:1.
- Gezahgne, A., Roux, J., Slippers, B., and Wingfield, M. J. (2004a). Identification of the causal agent of *Botryosphaeria* stem canker in Ethiopian *Eucalyptus* plantations. *South Afr. J. Bot.* 70, 241–248. doi: 10.1016/s0254-6299(15)30241-6
- Gezahgne, A., Roux, J., Thu, P. Q., and Wingfield, M. J. (2004b). *Coniothyrium* stem canker of *Eucalyptus*, new to Argentina and Vietnam. *South Afr. Forestry J.* 99, 578–588.
- Gezahgne, A., Roux, J., and Wingfield, M. J. (2003). Diseases of exotic *Eucalyptus* and *Pinus* species in Ethiopia plantations. *South Afr. Forestry J.* 99, 29–33.
- Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J., and Thompson, R. (2015). *ASReml Guide Release 4.1 Functional Specifications*. 4, 1 Edn. Hemel Hempstead: VSN International Ltd.
- Graziosi, I., Tembo, M., Kuate, J., and Muchugi, A. (2020). Pests and diseases of trees in Africa: a growing continental emergency. *Plants People Planet* 2, 14–28. doi: 10.1002/ppp3.31
- Guo, G., Zhao, F., Wang, Y., Zhang, Y., Du, L., and Su, S. (2014). Comparison of single-trait and multiple-trait genomic prediction models. *BCM Genet.* 15:30. doi: 10.1186/1471-2156-1115-1130
- Habier, D., Fernando, R. L., and Dekker, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177, 2389–2397. doi: 10.1534/genetics.107.081190
- Harun-or-Rashid, M., and Chung, Y. R. (2017). Induction of systemic resistance against insect herbivores in plants by beneficial soil microbes. *Front. Plant Sci.* 8:1816. doi: 10.3389/fpls.2017.01816
- Hayes, B. J., Visscher, P. M., and Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res.* 91, 47–60. doi: 10.1017/s0016672308009981
- Hill, W. G., and Weir, B. S. (2010). Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet. Resour.* 93, 47–64. doi: 10.1017/s0016672310000480
- Isik, F., Holland, J., and Maltecca, C. (2017). *Genetic Data Analysis for Plant and Animal Breeding*. Cham: Springer International Publishing. doi: 10.1007/978-3-319-55177-7_1
- Jactel, H., Desperz-Loustau, M.-L., Battisti, A., Brockerhoff, E., Santini, A., Stenlid, J., et al. (2020). Pathologists and entomologists must join forces against forest pest and pathogen invasions. *NeoBiota* 58, 107–127. doi: 10.3897/neobiota.58.54389
- Jimu, L., Chen, S., Wingfield, M. J., and Mwenje, E. (2015). Three genetic groups of the *Eucalyptus* stem canker pathogen *Teratosphaeria zuluensis* introduced into Africa from an unknown source. *Antonie van Leeuwenhoek* 109, 21–33. doi: 10.1111/efp.12095
- Kim, I.-K., Mendel, Z., Protasov, A., Blumberg, D., and Salle, J. L. (2008). Taxonomy, biology, and efficacy of two Australian parasitoids of the eucalyptus gall wasp, *Leptocybe invasa* Fisher & La Salle (Hymenoptera: Eulophidae: Tetrastichinae). *Zootaxa* 1910, 1–20. doi: 10.11646/zootaxa.1910.1.1
- Klapste, J., Dungey, H. S., Graham, N. J., and Telfer, E. J. (2020). Effect of trait's expression level on single-step genomic evaluation of resistance to *Dothistroma* needle blight. *BCM Plant Biol.* 20:205. doi: 10.1186/s12870-12020-02403-12876
- Klapste, J., Suontanna, M., Dungey, H. S., Telfer, E. J., Graham, N. J., Low, C. B., et al. (2018). Effect of hidden relatedness on single-step genetic evaluation in an advanced open-pollinated breeding program. *J. Heredity* 109, 802–810.
- Legarra, A., Aguilar, I., and Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92, 4656–4663. doi: 10.3168/jds.2009-2061
- Legarra, A., Christensen, O. F., Aguilar, I., and Misztal, I. (2014). Single Step, a general approach for genomic selection. *Livestock Sci.* 166, 54–65. doi: 10.1016/j.livsci.2014.04.029
- Li, Y., Klapste, J., Telfer, E., Wilcox, P., Graham, N., Macdonald, L., et al. (2019). Genomic selection for non-key traits in radiata pine when the documented pedigree is corrected using DNA marker information. *BCM Genomics* 20:1026. doi: 10.1186/s12864-12019-16420-12868
- Lourenco, D. A. L., Tsuruta, S., Fragomeni, B. O., Masuda, Y., Aguilar, I., Legarra, A., et al. (2015). Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *Am. Soc. Anim. Sci.* 93, 2653–2662. doi: 10.2527/jas.2014-8836
- Luo, J., Arnold, R., Lu, W., and Lin, Y. (2014). Genetic variation in *Eucalyptus camaldulensis* and *E. tereticornis* for early growth and susceptibility to the gall wasp *Leptocybe invasa* in China. *Euphytica* 196, 397–411. doi: 10.1007/s10681-013-1042-8
- Ma, P., Lund, M. S., Nielsen, U. S., Aamand, G. P., and Su, G. (2015). Single-step genomic model improved reliability and reduced the bias of genomic prediction in Danish Jersey. *J. Dairy Sci.* 98, 9026–9034. doi: 10.3168/jds.2015-9703
- Malan, F. S. (1993). The wood properties and qualities of three South African-grown eucalypt hybrids. *South Afr. Forestry J.* 167, 35–44. doi: 10.1080/00382167.1993.9629409
- Marsberg, A., Kemler, M., Jami, F., Nagel, J. H., Postma-Smidt, A., Naidoo, S., et al. (2017). *Botryosphaeria dothidea*: a latent pathogen of global importance to woody plant health. *Mol. Plant Pathol.* 18, 477–488. doi: 10.1111/mpp.12495
- Mendel, Z., Protasov, A., Fisher, N., and La Salle, J. (2004). Taxonomy and biology of *Leptocybe invasa* gen. & sp. n. (Hymenoptera: Eulophidae), an invasive gall inducer on *Eucalyptus*. *Aust. Entomol.* 43, 101–113. doi: 10.1111/j.1440-6055.2003.00393.x
- Mhoswa, L., Neill, O., Mphahlele, M. M., Oates, C. N., Payn, K. G., Slippers, B., et al. (2020). A Genome-wide association study for resistance to the insect pest *Leptocybe invasa* in *Eucalyptus grandis* reveals genomic regions and positional

- candidate defence genes. *Plant Cell Physiol.* 61, 1285–1296. doi: 10.1093/pcp/pcaa057
- Misztal, J., Aggrey, S. E., and Muir, W. M. (2013). Experiences with a single-step genome evaluation. *Poultry Sci.* 92, 2530–2534. doi: 10.3382/ps.2012-02739
- Misztal, J., Legarra, A., and Aguilar, I. (2009). Computing procedures for genetic evaluation including phenotypic, full and genomic information. *J. Dairy Sci.* 92, 4648–4655. doi: 10.3168/jds.2009-2064
- Mohali, S. A., Slippers, B., and Wingfield, M. J. (2007). Identification of *Botryosphaeriaceae* from *Eucalyptus*, *Acacia* and *Pinus* in Venezuela. *Fungal Divers.* 25, 103–125.
- Munoz, P. R., Resende, M. F. R. Jr., Huber, D. A., Quesada, T., Resende, M. D. V., Naele, D. B., et al. (2014). Genomic relationship matrix for correcting pedigree errors in breeding populations: impact on genetic parameters and genomic selection accuracy. *Crop Sci.* 54, 1115–1123. doi: 10.2135/cropsci2012.12.0673
- Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., et al. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356–375.
- Naidoo, S., Kulheim, C., Zwart, L., Mangwanda, R., Oates, C. N., Visser, E. A., et al. (2014). Uncovering the defence responses of *Eucalyptus* to pests and pathogens in the genomics age. *Tree Physiol.* 34, 931–943. doi: 10.1093/treephys/tpu075
- Naidoo, S., Slippers, B., Plett, J. M., Coles, D., and Oates, C. N. (2019). The road to resistance in forest trees. *Front. Plant Sci.* 10:273.
- Neser, S., Prinsloo, G. L., and Neser, O. C. (2007). The eucalypt leaf, twig and stem (sic) galling wasp, *Leptocybe invasa*, now in South Africa. *Plant Protection News* 72, 1–2.
- Nugnes, F., Gebiola, M., Monti, M. M., Gualtieri, L., Giorgini, M., Wang, J., et al. (2015). Genetic Diversity of the Invasive Gall Wasp *Leptocybe invasa* (Hymenoptera: Eulophidae) and of its *Rickettsia* Endosymbiont, and Associated Sex-Ratio Differences. *PLoS One* 10:e0124660. doi: 10.1371/journal.pone.0124660
- Nyeko, P. (2005). The cause, incidence and severity of a new gall damage on *Eucalyptus* species at Oruchinga refugee settlement in Mbarara district. Uganda. *Uganda J. Agric. Sci.* 11, 47–50.
- Nyeko, P., Mutitu, K. E., Otieno, B. O., Ngai, G. N., and Day, R. K. (2010). Variations in *Leptocybe invasa* (Hymenoptera: Eulophidae) population intensity and infestation on eucalyptus germplasms in Uganda and Kenya. *Int. J. Pest Manage.* 56, 137–144. doi: 10.1080/09670870903248835
- Perez, C. A., Altier, N., Simeto, S., Wingfield, M. J., Slippers, B., and Blanchette, R. A. (2008). *Botryosphaeriaceae* from *Eucalyptus* and native Myrtaceae in Uruguay. *Agrociencia* 12, 19–30.
- Perlack, R. D., Turhollow, W. L. L., Af, G. R. L., Stokes, B. J., and Erbach, D. C. (2005). *Biomass as a Feedstock for a Bioenergy and Bioproduct Industry: The Technical Feasibility of a Billion-ton Annual Supply*, ed. OAK RIDGE NATIONAL LAB TN (Washington, DC: Department of Energy).
- Prabhu, J. S. T. (2010). Susceptibility of *Eucalyptus* species and clones to gall wasp, *Leptocybe invasa* Fisher and La Salle (Eulophidae: Hymenoptera) in Karnataka. *Karnataka J. Agric. Sci.* 23, 220–221.
- Ratcliffe, B., El-Dien, O. G., Cappa, E. P., Porth, I., Klápšte, J., Chen, C., et al. (2017). Single-step BLUP with varying genotyping effort in open-pollinated *Picea glauca*. *Genes Genomics Genet.* 7, 935–942. doi: 10.1534/g3.116.037895
- Raymond, C. A. (2002). Genetics of *Eucalyptus* wood properties. *Annu. Forestry Sci.* 59, 525–531. doi: 10.1051/forest:2002037
- Rodas, C. A., Slippers, B., Gryzenhout, M., and Wingfield, M. J. (2009). *Botryosphaeriaceae* associated with *Eucalyptus* canker diseases in Colombia. *Forest Pathol.* 39, 110–123.
- Roux, J., Coutinho, T. A., Wingfield, M. J., and Bouillet, J.-P. (2000). Diseases of plantation *Eucalyptus* in the Republic of Congo. *South Afr. J. Sci.* 96, 454–456.
- Sangtongpraow, B., Charensom, K., and Siripatanadilok, S. (2011). Longevity, fecundity and development time of *Eucalyptus* gall wasp, *Leptocybe invasa* Fisher and La Salle (Hymenoptera: Eulophidae) in Kanchanaburi province, Thailand. *Thai J. Agric. Sci.* 44, 155–163.
- Silva, P. H. M., Miranda, A. C., Moraes, M. L. T., Furtado, E. L., Stape, J. L., Alvares, C. A., et al. (2013). Selecting for rust (*Puccinia psidii*) resistance in *Eucalyptus grandis* in Sao Paulo State, Brazil. *Forest Ecol. Manage.* 303, 91–97. doi: 10.1016/j.foreco.2013.04.002
- Silva-Junior, O. B., Faria, D. A., and Grattapaglia, D. (2015). A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytologist* 206, 1527–1540. doi: 10.1111/nph.13322
- Slippers, B., Burgess, T., Pavlic, D., Ahumada, R., Maleme, H., Mohali, S., et al. (2009). A diverse assemblage of *Botryosphaeriaceae* infects *Eucalyptus* in native and non-native environments. *Southern Forests* 71, 101–110. doi: 10.2989/sf.2009.71.2.3.818
- Smith, H., Kemp, G. H. J., and Wingfield, M. J. (1994). Canker and dieback of *Eucalyptus* in South Africa caused by *Botryosphaeria dothidea*. *Plant Pathol.* 43, 1031–1034. doi: 10.1111/j.1365-3059.1994.tb01653.x
- Smith, H., Wingfield, M. J., and Petrini, O. (1996). *Botryosphaeria dothidea* endophytic *Eucalyptus grandis* and *Eucalyptus nitens* in South Africa. *Forest Ecol. Manage.* 89, 189–195. doi: 10.1016/s0378-1127(96)03847-9
- Stafford, W., De Lange, W., Nahman, A., Chunilall, V., Lekha, P., Andrew, J., et al. (2020). Forestry biorefineries. *Renewable Energy* 154, 461–475.
- Thu, P. Q., Dell, B., and Burgess, T. I. (2009). Susceptibility of 18 Eucalypt species to the gall wasp *Leptocybe invasa* in the nursery and young plantations in Vietnam. *ScienceAsia* 35, 113–117.
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980
- Viggiani, G., Laudonia, S., and Bernardo, U. (2000). The increase of insect pests in *Eucalyptus*. *Informatore Agrario* 58, 86–87.
- Vivas, M., Kemler, M., Mphahlele, M. M., Wingfields, M. J., and Slippers, B. (2017). Maternal effects on phenotype, resistance and the structuring of fungal communities in *Eucalyptus grandis*. *Environ. Exp. Bot.* 140, 120–127. doi: 10.1016/j.envexpbot.2017.06.002
- Wani, S. H., Kumar, V., Shriram, V., and Sah, S. K. (2016). Phytohormones and their metabolic engineering for abiotic stress tolerance in crop plants. *Crop J.* 4, 162–176. doi: 10.1016/j.cj.2016.01.010
- Wiley, J., and Skelly, P. (2008). Pest Alert: a *Eucalyptus* pest, *Leptocybe invasa* Fisher and LaSalle (Hymenoptera: Eulophidae), genus and species new to Florida and North America. *Florida Dept. Agric. Consum. Serv.* 23, 359–374.
- Wimmer, V., Albrecht, T., Auinger, H.-J., and Schon, C.-C. (2012). synbreed: a framework for the analysis of genomic predictions data using R. *Bioinformatics* 28, 2086–2087. doi: 10.1093/bioinformatics/bts335
- Wingfield, M. J., Brocherhoff, E. G., Wingfield, B. D., and Slippers, B. (2015). Planted forest health: The need for a global strategy. *Science* 349, 832–836. doi: 10.1126/science.aac6674
- Wingfield, M. J., Crous, P. W., and Coutinho, T. A. (1996). A serious canker disease of *Eucalyptus* in South Africa caused by a new species of *Coniothyrium*. *Mycopathologia* 136, 139–145. doi: 10.1007/bf00438919
- Zhang, M., Zhou, C., Song, Z., Weng, Q., Li, M., Ji, H., et al. (2018). The first identification of genomic loci in plants associated with resistance to galling insects: a case study in *Eucalyptus* L'Hér. (Myrtaceae). *Sci. Rep.* 8:2319. doi: 10.1038/s41598-018-20780-41599
- Zheng, X.-L., Huang, Z.-Y., Dong, D., Guo, C.-H., Li, J., Yang, Z.-D., et al. (2016). Parasitoids of the eucalyptus gall wasp *Leptocybe invasa* (Hymenoptera: Eulophidae) in China. *Parasite* 23:58. doi: 10.1051/parasite/2016071
- Zhu, F.-L., Ren, S.-X., Giu, B.-L., Huang, Z., and Peng, Z.-Q. (2012). The abundance and population dynamics of *Leptocybe invasa* (Hymenoptera: Eulophidae) galls on *Eucalyptus* spp. in China. *J. Integr. Agric.* 11, 2116–2123. doi: 10.1016/s2095-3119(12)60470-5

Conflict of Interest: MM was employed by the company Mondi South Africa (Pty) Ltd. The authors declare that this study received funding from Mondi South Africa (Pty) Ltd. The funder was not involved in the study design, collection, analysis, and interpretation of data, the writing of this article, or the decision to submit it for publication.

Copyright © 2021 Mphahlele, Isik, Hodge and Myburg. 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.



Dissecting Quantitative Trait Loci for Spot Blotch Resistance in South Asia Using Two Wheat Recombinant Inbred Line Populations

Chandan Roy¹, Navin C. Gahtyari², Xinyao He³, Vinod K. Mishra⁴, Ramesh Chand⁵, Arun K. Joshi⁶ and Pawan K. Singh^{3*}

¹ Department of Plant Breeding and Genetics, Bihar Agricultural University, Sabour, India, ² ICAR–Vivekanand Parvatiya Krishi Anushandhan Sansthan, Almora, India, ³ International Maize and Wheat Improvement Center (CIMMYT), Texcoco, Mexico, ⁴ Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, India, ⁵ Department of Mycology and Plant Pathology, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, India, ⁶ CIMMYT-India/Borlaug Institute for South Asia, New Delhi, India

OPEN ACCESS

Edited by:

Anna Maria Mastrangelo,
Research Centre for Cereal
and Industrial Crops, Council
for Agricultural and Economics
Research (CREA), Italy

Reviewed by:

Bernardo Ordas,
Consejo Superior de Investigaciones
Científicas (CSIC), Spain
Weiya Xue,
Pennsylvania State University (PSU),
United States

*Correspondence:

Pawan K. Singh
pk.singh@cgiar.org

Specialty section:

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

Received: 14 December 2020

Accepted: 26 January 2021

Published: 04 March 2021

Citation:

Roy C, Gahtyari NC, He X,
Mishra VK, Chand R, Joshi AK and
Singh PK (2021) Dissecting
Quantitative Trait Loci for Spot Blotch
Resistance in South Asia Using Two
Wheat Recombinant Inbred Line
Populations.
Front. Plant Sci. 12:641324.
doi: 10.3389/fpls.2021.641324

Spot blotch (SB) disease causes significant yield loss in wheat production in the warm and humid regions of the eastern Gangetic plains (EGP) of South Asia (SA). Most of the cultivated varieties in the eastern part of SA are affected by SB under favorable climatic conditions. To understand the nature of SB resistance and map the underlying resistant loci effective in SA, two bi-parental mapping populations were evaluated for 3 years, i.e., 2013–2015 for the BARTAI × CIANO T79 population (denoted as BC) and 2014–2016 for the CASCABEL × CIANO T79 population (CC), at Varanasi, Uttar Pradesh, India. DArTSeq genotyping-by-sequencing (GBS) platform was used for genotyping of the populations. Distribution of disease reaction of genotypes in both populations was continuous, revealing the quantitative nature of resistance. Significant “genotype,” “year,” and “genotype × year” interactions for SB were observed. Linkage map with the genome coverage of 8,598.3 and 9,024.7 cM in the BC and CC population, respectively, was observed. Two quantitative trait loci (QTLs) were detected on chromosomes 1A and 4D in the BC population with an average contribution of 4.01 and 12.23% of the total phenotypic variation (PV), respectively. Seven stable QTLs were detected on chromosomes 1B, 5A, 5B, 6A, 7A, and 7B in the CC population explaining 2.89–10.32% of PV and collectively 39.91% of the total PV. The QTL detected at the distal end of 5A chromosome contributed 10.32% of the total PV. The QTLs on 6A and 7B in CC could be new, and the one on 5B may represent the *Sb2* gene. These QTLs could be used in SB resistance cultivar development for SA.

Keywords: *Bipolaris sorokiniana*, SNPs, bi-parental mapping, DArTSeq, wheat QTLs for SB resistance

INTRODUCTION

Spot blotch (SB) disease caused by *Cochliobolus sativus* (Ito and Kuribayashi) Drechsler ex Dastur [anamorph *Bipolaris sorokiniana* (Sacc.) Shoemaker] is considered a significant disease of wheat (*Triticum aestivum* L.) in South Asia (SA) (Gupta et al., 2018). High temperature and humidity favor the disease development in the warmer wheat growing areas of the eastern Gangetic plains

(EGP), particularly in Bangladesh (Siddique et al., 2006), Nepal (Sharma et al., 2007b), and eastern India (Joshi et al., 2007a). The long-established practice of rice–wheat cropping system in the EGP delays the sowing of wheat crop that provides a congenial humid and warm environment for the SB development in the later stages of crop growth. Average yield loss due to SB ranges from 15 to 20%, but under favorable environment, up to 87% yield loss has been observed in susceptible genotypes (Hetzler et al., 1991). Delayed seeding of wheat in the EGP resulted on an average loss of 30% yield due to complex foliar blights, especially SB (Duveiller et al., 2005). Trait association analysis revealed that days to heading (DH) and plant height (PH) often showed a negative correlation with SB severity (Singh et al., 2015). Several attempts have been made including cultural practices and chemical application to control SB, but none of them was completely successful. Integrated disease management using host resistance, chemical control, and cultural practices is considered most effective in managing the disease (Joshi and Chand, 2002).

Except for a recent attempt (Kumar et al., 2019), no host immunity has been observed for SB, and the best released cultivars are only partially resistant. SB resistance is under polygenic control, with quantitative trait loci (QTL) of various phenotypic effects; hence, the progress of cultivar development is relatively slow. Genetic studies for SB resistance have identified multiple QTLs, of which four with major effects have been nominated, i.e., *Sb1* through *Sb4*. The *Sb1* is located on chromosome 7DS flanked by the markers *Xgwm1220* and *Xgwm295*, being co-located with the leaf rust resistance locus *Lr34* having pleiotropic effects on resistance to yellow rust (*Yr57*), powdery mildew (*Pm8*), and leaf tip necrosis (*Ltn+*) (Lillemo et al., 2013). The *Sb2* (*Qsb.bhu-5B*) has been mapped on chromosome 5BL flanked by the simple sequence repeat (SSR) markers *Xgwm639* and *Xgwm1043* (Kumar et al., 2015). The third gene, *Sb3*, was mapped on chromosome 3BS (Lu et al., 2016), being in the same region where two previously reported QTLs *Qsb.bhu 3B* and *Qsb.cim 3B* reside. Recently, the *Sb4* gene has been mapped on the long arm of chromosome 4B, where 21 putative genes were predicted (Zhang et al., 2020). QTLs with minor effects are also important for SB resistance since stacking such QTLs significantly reduced SB severity (Singh et al., 2018). Multiple minor QTLs have been mapped on 1A, 1B, 1D, 2B, 2D, 3A, 3B, 4A, 5A, 5B, 6A, and 7A (Gurung et al., 2014; Zhu et al., 2014; Singh et al., 2016; Bainsla et al., 2020).

Germplasm development for SB resistance started in the 1980s, which led to the identification of several wheat genotypes with variable resistance like Saar, Yangmai 6, Shanghai 4, M3, Chirya 1, Chirya 3, Chirya 7, and SYN1 (Ibeagha et al., 2005). Looking at the growing incidence of SB in SA, CIMMYT developed a special nursery in 2009 for SA named CSISA-SB, under the Cereal System Initiative for South Asia (CSISA) project. The purpose was to share CIMMYT breeding lines with SB resistance and good agronomic performance with the researchers of other countries and to test the nursery over various locations. This nursery was renamed Helminthosporium Leaf Blight Screening Nursery (HLBSN) in 2015 and distributed

beyond SA to South American and African countries like Brazil, Bolivia, Paraguay, and Zambia, where SB is of major concern. The SB screening platform of CIMMYT in Mexico is located at Agua Fria, where the climate is similar to SA, providing strong support in the selection of SB-resistant genotypes for SA (Singh et al., 2015).

In SA, the Varanasi center of India has been identified as one of the most suitable sites for the evaluation of SB; it has a close similarity with the climatic conditions of Bhairahawa and Rampur of Nepal (Joshi et al., 2007a). Previously, four bi-parental mapping populations were evaluated at Agua Fria, Mexico, and their underlying QTLs have been identified (Singh et al., 2018; He et al., 2020). In the present study, we evaluated two of those four mapping populations at Varanasi, India, to determine the resistant QTLs effective under the SA environment.

MATERIALS AND METHODS

Plant Materials

Two SB-resistant lines BARTAI (BABAX/LR 42//BABAX/3/ERA F 2000) and CASCABEL (SOKOLL/W15.92/WEEBILL1) identified in the previous experiments were crossed with a common susceptible parent CIANO T79 (BUCKY/(SIB)MAYA-74/4/BLUEBIRD//HD-832.5.5/OLESEN/3/CIANO-67/PENJAMO-62) to develop two bi-parental mapping populations (Singh et al., 2018; He et al., 2020). Recombinant inbred lines (RILs) were generated following the single seed descent method from F₂ generation of the cross BARTAI × CIANO T79 (BC population) and CASCABEL × CIANO T79 (CC population) at CIMMYT, Mexico. Field experiments were conducted using a total of 231 RILs of BC and 226 RILs of CC in F_{2:7} generation along with the parents constituting the populations, and genotypes Chirya 3 and Sonalika were included as resistant and susceptible check, respectively.

Field Experiments

Field evaluation was carried out at the experimental station of Banaras Hindu University (BHU, 25.2°N, 83.0°E), Varanasi, India, in the years 2012–2013 (denoted as 2013), 2013–2014 (2014), and 2014–2015 (2015) for the BC population, and in the years 2013–2014 (2014), 2014–2015 (2015), and 2015–2016 (2016) for the CC population. Sowing was done in December, under late sown conditions to expose the crop to high temperature and humidity at the later stage of crop growth, which favors SB disease development. The experiments were conducted in a randomized complete block design with two replications, where each entry was sown in 2-m double rows spaced 25 cm apart, with a plant-to-plant distance of 5 cm.

Inoculation Method and SB Assessment

The pure culture of *B. sorokiniana* (isolate HD 3069/MCC 1572) was maintained using potato dextrose agar (PDA) medium (Chand et al., 2003). The pathogen was mass multiplied on previously soaked and autoclaved sorghum grains, which was kept under room temperature for at least 6 weeks. Spore

suspension culture was prepared at a concentration of 1×10^4 spores ml^{-1} . To create an artificial epiphytotic condition, the spore suspension was inoculated at the heading stage [Zadok's growth stage (GS) 55] in the evening time. Light irrigation was given after inoculation to maintain high humidity for disease development.

Disease scoring was done for three subsequent growth stages at the beginning of anthesis (GS 63), after completion of anthesis (GS 69), and late milking (GS 77) using a double-digit (00–99) scale as prescribed by Saari and Prescott (1975). The first digit (D1) measured disease progress in PH and the second digit (D2) measured the disease severity in terms of the proportion of infected leaf area. The percentage of disease severity for each score was measured as:

$$\text{Severity}(\%) = (D1/9) \times (D2/9) \times 100$$

Area under disease progress curve (AUDPC) was calculated using percent of severity estimations corresponding to disease rating as:

$$\text{AUDPC} = \sum_{i=1}^n [(Y_i + Y_{i+1})/2] \times \{t_{i+1} - t_i\}$$

where

Y_i = disease level at the time t_i .

$Y_{(i+1)}$ = disease level at time t_{i+1} .

$t_{i+1} - t_i$ = time difference in days between two disease scores.

n = number of readings.

Area under disease progress curve average of two replications in a single year and mean AUDPC across all 3 years were used for QTL analysis. DH and PH were also measured to determine their association with SB.

Statistical Analysis

Analysis of variance (ANOVA) and Pearson correlation coefficients were calculated using statistical software OPSTAT¹. Marker-based narrow sense heritability was calculated with the “heritability” package of R (Kruijer et al., 2015).

Genotyping and Linkage Analysis

Genomic DNA was isolated using the cetyltrimethylammonium bromide (CTAB) method from each entry including the parental lines of respective populations. Genotyping was carried out using the DArTSeq genotyping-by-sequencing (GBS) platform (Li et al., 2015) at the Genetic Analysis Service for Agriculture (SAGA) in Guadalajara, Mexico. Several gene-based markers and D-genome-specific single-nucleotide polymorphism (SNP) markers using “Kompetitive allele-specific PCR” (KASP) were also used. QTL analysis was carried out using an integrated software package ICIMapping version 4.1 (Meng et al., 2015). Monomorphic markers, markers with missing value >20%, and minor allele frequency <30% were removed from QTL analysis. Chromosome anchoring was done for each marker as per the GBS map described by Li et al. (2015). Linkage groups (LGs) were constructed using the MAP function in the ICIMapping

software version 4.1, with the LOD threshold set at 15 and the rest parameters at default.

Quantitative trait loci mapping was performed using the BIP function of ICIMapping, where interval mapping was first carried out to identify significant QTLs and after that inclusive composite interval mapping was performed to identify more robust QTLs. QTL mapping was also carried out after adjusting for DH and PH. Adjusted mean was calculated by the software Multi-Environment Trial Analysis with R (META-R) version 6.0 using DH and PH as cofactors. A QTL was considered significant when it exceeded the LOD threshold of 3.4 (1,000 permutations at $\alpha = 0.05$) for BC and LOD of 3.6 for CC populations in at least one environment. However, QTL with an LOD value of 2.5 or above appearing in more than one environment was also considered as significant. To draw the LGs and LOD curve, software MapChart v. 2.3 (Voorrips, 2002) was used.

RESULTS

Phenotyping for SB Resistance

Significant genetic variation was observed for SB among the genotypes in both the BC and CC populations. Effects of climatic fluctuations across years on SB development were revealed by significant variation in “year” and “genotype \times year” interaction effects; however, for the CC population, the latter effect was non-significant (Table 1). Disease pressure was maximum in the year 2014 and least in 2015 for both populations. A similar trend was also observed for Sonalika and Chirya 3, the susceptible and resistant checks, respectively. Continuous distribution of genotypes for SB resistance in different years and their mean were observed (Figure 1). Transgressive segregants for resistance and susceptibility were obtained in both the populations. Twenty-three resistant transgressive segregants were found in the BC population, out of which seven genotypes performed better than the resistant check Chirya 3, whereas in the CC population, 55 genotypes showed higher

TABLE 1 | Analysis of variance and heritability estimates of spot blotch resistance in BARTAI \times CIANOT79 (BC) and CASCABEL \times CIANOT79 (CC) populations.

| Source of variation | DF | Mean squares | F calculated | Significance | Heritability |
|------------------------|-----|--------------|--------------|--------------|--------------|
| BC population | | | | | |
| Year | 2 | 13,488.82 | 5,748.176 | <0.0001 | |
| Rep (year) | 3 | 2.34 | 2.46 | NS | |
| Genotype | 230 | 58.91 | 2.784 | <0.0001 | 0.61 |
| Year \times genotype | 460 | 21.15 | 21.162 | <0.0001 | |
| Pooled error | 690 | 0.95 | | | |
| CC Population | | | | | |
| Year | 2 | 103.37 | 58.033 | <0.01 | |
| Rep (year) | 3 | 1.78 | 1.78 | NS | |
| Genotype | 225 | 28.58 | 126.553 | 0.0001 | 0.73 |
| Year \times genotype | 450 | 0.22 | 0.226 | NS | |
| Pooled error | 675 | 1.00 | | | |

NS, non-significant.

¹<http://14.139.232.166/opstat/>

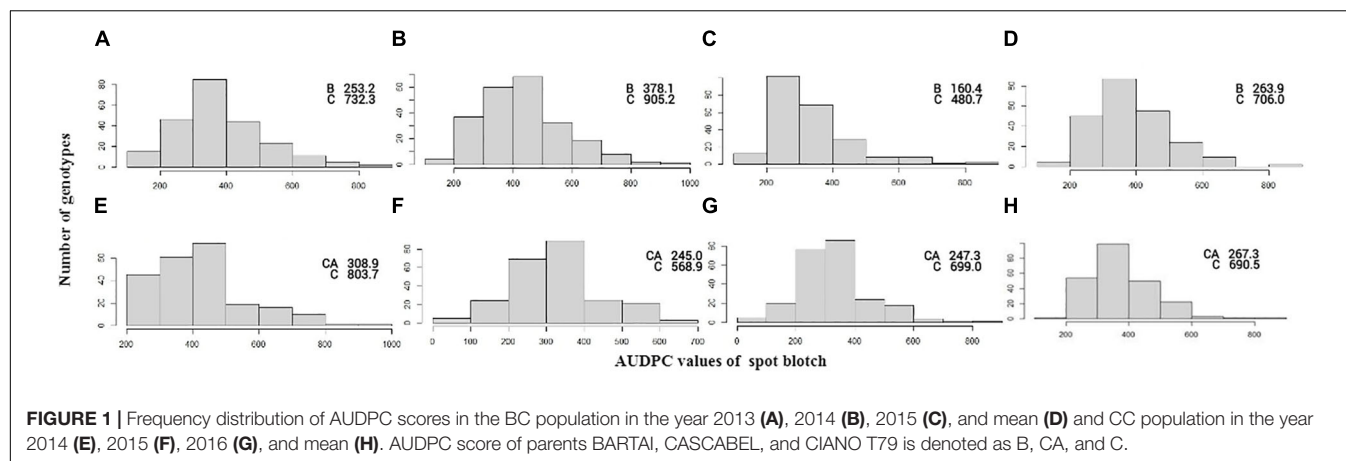


FIGURE 1 | Frequency distribution of AUDPC scores in the BC population in the year 2013 (A), 2014 (B), 2015 (C), and mean (D) and CC population in the year 2014 (E), 2015 (F), 2016 (G), and mean (H). AUDPC score of parents BARTAI, CASCABEL, and CIANO T79 is denoted as B, CA, and C.

resistant than CASCABEL, out of which 10 genotypes were better than Chirya 3.

Moderate heritability estimates for SB in BC (0.61) and CC (0.73) were recorded (Table 1). DH and PH were mostly negatively correlated with SB. PH was found to be more closely associated with SB than DH as exhibited in the significantly negative association across all the environments in both populations (Table 2).

Genotyping and Linkage Analysis

Out of 18,000 GBS markers scored in both populations, 3,174 and 3,197 high-quality non-redundant markers in BC and CC populations, respectively, were screened out for linkage analysis and QTL mapping (Supplementary Tables 1, 2). Both populations contained 21 large LGs representing all the 21 wheat chromosomes, as well as a few fragmented LGs that were not used in subsequent analysis. The linkage map of the BC population covered 8,598.3 cM with an average distance of 2.71 cM between markers, while in the CC population, 9024.21 cM was covered with an average distance of 2.82 cM between markers. All chromosomes were in good coverage with the least length of 199.22 cM for 1D in the BC population and 223.69 cM for 4D in the CC population. The coverage for chromosomes of A and B sub-genome was better than that of D genome in both populations (Supplementary Tables 1, 2).

TABLE 2 | Pearson correlation coefficient analysis of spot blotch resistance with days to heading and plant height in BARTAI × CIANO T79 (BC) and CASCABEL × CIANO T79 (CC) populations.

| Year | 2013 | 2014 | 2015 |
|----------------------|----------------------|---------------------|----------------------|
| BC population | | | |
| Days to heading | −0.430** | 0.033 ^{NS} | −0.111 ^{NS} |
| Plant height | −0.432** | −0.344** | −0.230** |
| CC population | | | |
| Year | 2014 | 2015 | 2016 |
| Days to heading | −0.096 ^{NS} | −0.194** | −0.121 ^{NS} |
| Plant height | −0.296** | −0.299** | −0.334** |

** $P < 0.001$; ^{NS} non-significant.

QTL Identification for SB Resistance

The QTL with the largest phenotypic variation explained (PVE) in the BC population was detected on chromosome 4D, with a mean PVE of 12.23% (Table 3). This QTL was found to be associated with PH. Another QTL was detected on chromosome 1A with the mean phenotypic effect of 4%. Two additional QTLs with minor effects were detected on 4D in 2013 and 5B in 2015 (Table 3 and Figure 2). All the resistance alleles of the QTLs in BC population were contributed by the resistant parent BARTAI. When PH was used as a covariate, the effect of the QTL reduced; in addition, few other QTLs on 4B, 5B, and 6B chromosomes were detected (Supplementary Table 3).

Seven QTLs were detected in the CC population, altogether explaining 39.91% of phenotypic variation (PV). The QTLs were detected on six chromosomes (1B, 5A, 5B, 6A, 7A, and 7B) with the mean PVE ranging from 2.89 to 10.32%. The resistance alleles of QTLs on 1B, 5A (proximal), 7A, and 7B were contributed by CIANO T79, whereas those of QTLs on 5A (distal), 5B, and 6A were contributed by CASCABEL. Out of 55 resistant transgressive segregants, 40 were carrying at least one QTL contributed by CIANO T79. The major QTL was detected on the distal end of chromosome 5A with a PV ranging from 8.93 to 12.62% in 2015 and 2014, respectively, with a mean PV of 10.32%. The effect of QTLs in the reduction of AUDPC appeared to be additive in nature (Figure 3).

DISCUSSION

Due to global warming and climate change, wheat production is predicted to be adversely affected (Rosyara et al., 2010). Higher temperature combined with rains during the grain filling stages increases the chances of SB in EGP, resulting in significant yield losses. Since the vulnerability of the wheat crop to SB increases when temperature exceeds 26°C at post-anthesis stage (Chaurasia et al., 2000), sowing in this study was delayed, exposing the populations under higher disease pressure. Joishi et al. (2002) reported that the resistance is independent of plant growth stage as there is appearance of substantial proportion of resistance in tall and dwarf progenies

TABLE 3 | Quantitative trait loci (QTLs) identified for SB in the BARTAI × CIANO T79 (BC) and CASCABEL × CIANO T79 (CC) populations and their associated QTLs in literature.

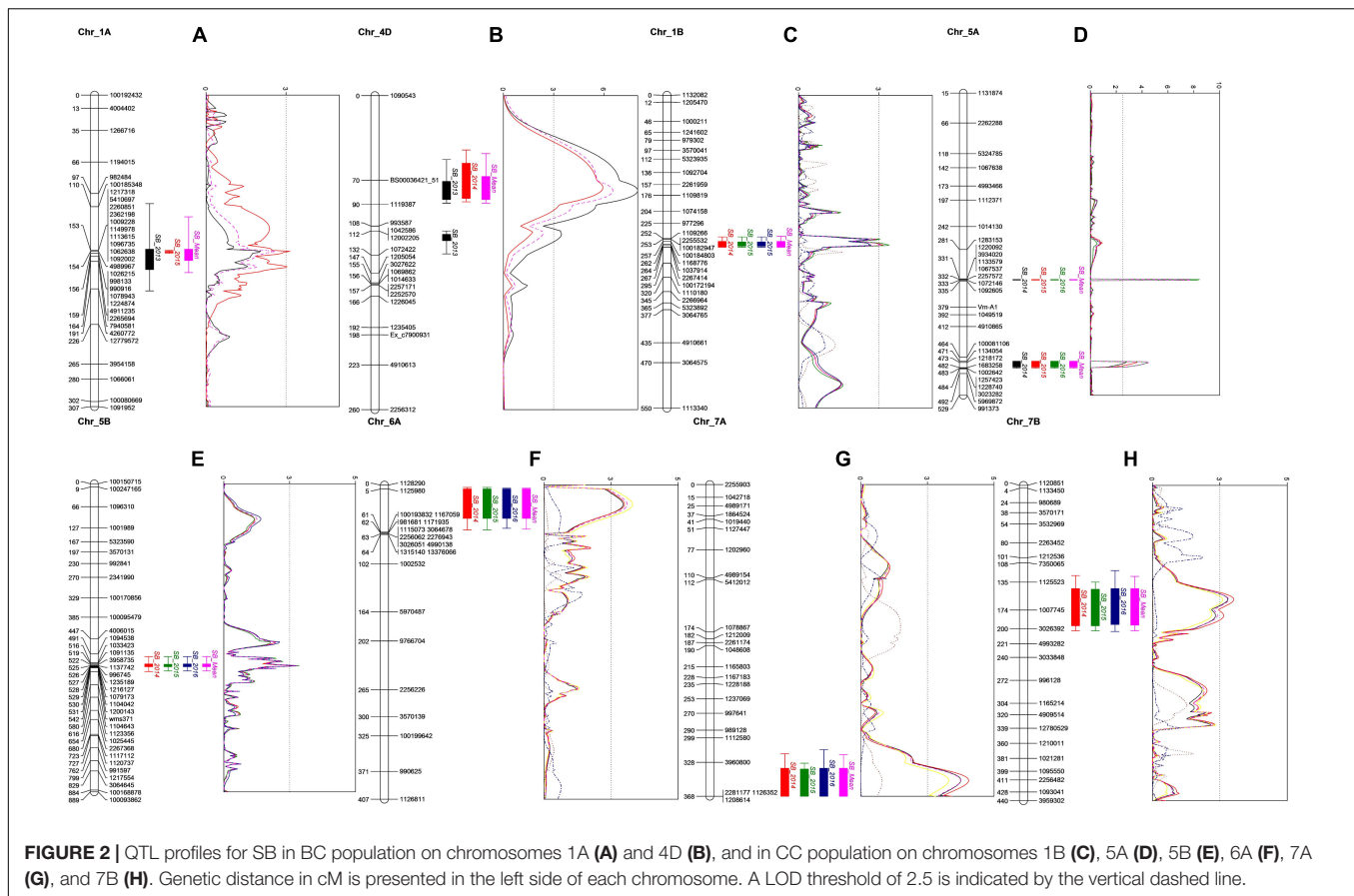
| Chromosome | Position (cM) | Left marker ^b | Right marker | LOD ^c | 2013 ^d | 2014 | 2015 | Mean | Source of resistance | Associated QTLs in literature |
|--|---------------|--------------------------|--------------|------------------|-------------------|------|-------|-------|----------------------|---|
| BARTAI × CIANO T 79 (BC) | | | | | | | | | | |
| 1A | 153.79–155.61 | 4989967 | 1026215 | 2.86 | 2.31 | – | 3.97 | 4.01 | BARTAI | Zhu et al., 2014 |
| 4D ^a | 70.49–90.31 | BS00036421_51 | 1119387 | 6.40 | 10.45 | 9.53 | – | 12.23 | BARTAI | Singh et al., 2018; He et al., 2020 |
| 4D | 112.48–131.93 | 12002205 | 1072422 | 3.46 | 4.49 | – | – | – | BARTAI | |
| 5B | 135.84–138.59 | 9724385 | 2267710 | 2.59 | – | – | 4.75 | – | BARTAI | Jamil et al., 2018; He et al., 2020 |
| Percentage of accumulated phenotypic variation | | | | | | | | 16.24 | | |
| CASCABEL × CIANO T 79 (CC) | | | | | | | | | | |
| 1B | 261.53–263.82 | 1168776 | 1037914 | 3.11 | 3.11 | 2.90 | 2.79 | 2.89 | CIANO T79 | Singh et al., 2018; Bainsla et al., 2020; He et al., 2020 |
| 5A ^a | 331.49–332.06 | 1067537 | 2257572 | 7.72 | 12.62 | 8.93 | 11.26 | 10.32 | CASCABEL | Ayana et al., 2018; Singh et al., 2018; Bainsla et al., 2020; He et al., 2020 |
| 5A | 472.56–481.59 | 1218172 | 1683258 | 4.51 | 4.59 | 4.13 | 5.71 | 5.89 | CIANO T79 | |
| 5B | 522.08–525.15 | 3958735 | 1137742 | 2.74 | 5.32 | 5.77 | 4.02 | 5.56 | CASCABEL | Kumar et al., 2009, 2010, 2015; Jamil et al., 2018; He et al., 2020 |
| 6A | 5.08–61.14 | 1125980 | 100193832 | 3.10 | 5.79 | 5.24 | 6.48 | 5.70 | CASCABEL | |
| 7A | 328.34–367.52 | 1126352 | 1208614 | 3.74 | 5.45 | 5.32 | 5.19 | 5.27 | CIANO T79 | |
| 7B | 134.52–173.55 | 1125523 | 1007745 | 2.82 | 4.56 | 4.18 | 4.31 | 4.28 | CIANO T79 | Singh et al., 2016; Ayana et al., 2018; Singh et al., 2018 |
| Percentage of accumulated phenotypic variation | | | | | | | | 39.91 | | |

^aNumber used to distinguish QTLs exceeded the LOD threshold of 3.4 for BC and 3.6 for CC population.

^bSequence information of the markers is available in **Supplementary Table 4**.

^cLOD values of QTLs in the mean population were used.

^dPercentage of phenotypic variation explained (PVE) is provided; QTLs in bold remained significant after the ICIM algorithm.

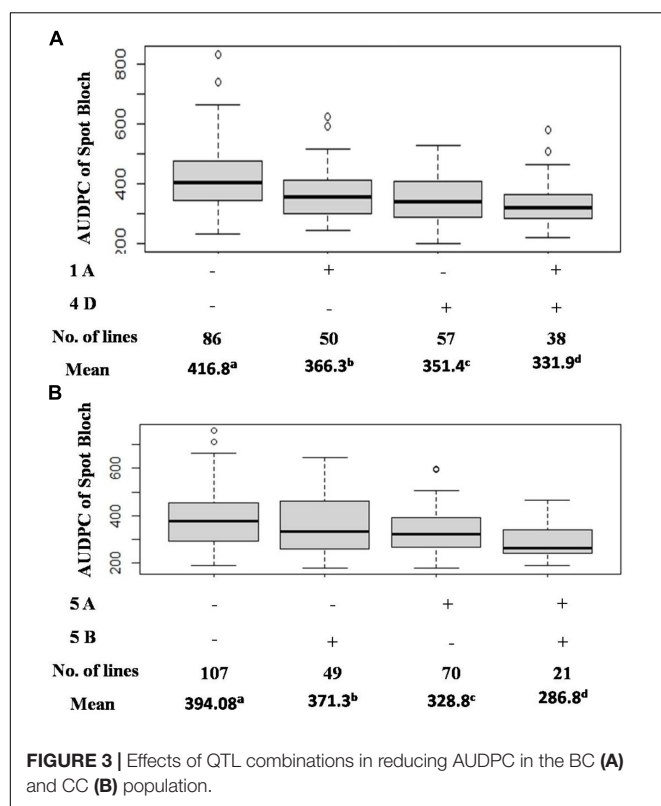


obtained from resistant tall \times susceptible dwarf and in early and late progenies from resistant tall \times susceptible early cross, respectively. However, to avoid the influence of growth stages on the disease appearance, scoring was done at different growth stages. The present observation on the negative correlation of DH and PH with SB has shown similarity with the previous reports from Agua Fria (Singh et al., 2018) and SA (Singh et al., 2016). Negative association of DH and PH with SB implies the selection of late and tall genotypes for better SB resistance, but such cultivars are not suitable for high rainfall and warmer regions like EGP of SA (Joshi et al., 2007b). Fortunately, this association has been broken, and several early maturing SB resistance cultivars have been identified for warm climatic conditions of SA (Sharma et al., 2004; Joshi et al., 2007b).

All experiments in the current study exhibited a typical quantitative inheritance of SB resistance with strong genotype \times environment ($G \times E$) interactions, which has also been reported in earlier studies constituting Indian germplasm, CIMMYT derivatives, and Afghan landraces (Kumar et al., 2016; Singh et al., 2018; Bainsla et al., 2020). Strong $G \times E$ interaction has always been a concern for plant breeders as it influences varietal adaptation across the environments. QTL \times environment interaction, a component of $G \times E$ interaction, affects the efficiency of marker-assisted selection. Identification of QTLs across the location and year helps breeders in design and implementation of breeding strategies

for the improvement of complex traits for adaptation in specific or mega environment. Veldboom and Lee (1996) reported that QTLs identified in the mean and across the environment are of major importance. In our study, two stable QTLs on chromosomes 1A and 4D were detected in the BC population, with the latter explaining major PVs. In a previous report, when the same population was evaluated at CIMMYT's Agua Fria station, a QTL on 4D chromosome was mapped at the same chromosomal region but was significant only in 1 year, having a PVE of 3.6% (Singh et al., 2018), implying a stronger influence of PH on SB resistance in SA. The QTL on 1A was not found in Singh et al. (2018) and thus might be specific only to SA environments. In a previous study, Zhu et al. (2014) identified a QTL on 1A, being close to the QTL identified in the present study.

A QTL on chromosome 5A delimited by the flanking markers 1067537–2257572 was significant in all 3 years explaining major PV in CC population. The same QTL was identified in the previous studies at Agua Fria but with higher phenotypic effects that reduced significantly when adjusted for PH and DH (Singh et al., 2018; He et al., 2020). In the current study, the QTL was identified 47.0 cM distance away from the vernalization locus *Vrn-A1*; also, the effect of QTL remained significant in all the 3 years when DH and PH were used as covariates. The allele *vrn-A1* responsible for late maturity was associated with SB resistance, escaping the disease due to delayed phenology.



However, the possibility of the presence of a SB resistance QTL in close proximity of *Vrn-A1* cannot be excluded (Singh et al., 2018; He et al., 2020) due to remnant effects of the QTL after adjusting for DH. Similarly, Bainsla et al. (2020) reported a marker 0.8 Mb away from *Vrn-A1* that is responsible for SB resistance. Likewise, Zhu et al. (2014) mapped a SB resistance QTL *Qsb.cim-5A* at 30.3 cM away from *Vrn-A1*. The strong marker trait association would be useful for the selection of SB resistance QTL based on *Vrn-A1*.

The QTL identified on 5B chromosome in the CC population was also reported earlier when evaluated at Agua Fria (He et al., 2020). However, in the present study, the contribution of this QTL was lower than the previous report, ranging from 4.02 to 5.77% in different years of evaluation. Comparing the QTL position using the sequences of flanking markers through BLAST to the IWGSC RefSeq v1.0 genome sequence of Chinese Spring (CS), the position of this QTL coincides with the previously identified *Sb2* gene (He et al., 2020). The presence of *Sb2* was reported in resistant genotypes Yangmai 6 (Kumar et al., 2009), Ning8201 (Kumar et al., 2010) and CASCABEL (He et al., 2020). Additionally, this gene has been detected in an Afghan population (Bainsla et al., 2020), CIMMYT germplasm (Jamil et al., 2018), and a diverse germplasm panel with global origin (Gurung et al., 2014), suggesting that *Sb2* has been selected by breeders of different continents due to its positive effects on SB resistance. Recently, *Tsn1* on 5BL was identified as a sensitivity gene for the pathogen carrying the corresponding virulent gene *ToxA* (Friesen et al., 2018). It was suggested that *Tsn1* gene is the susceptibility gene of *Sb2* (Friesen et al., 2018), but He et al. (2020)

proposed that the two genes might be different. A recent study of Indian *B. sorokiniana* population indicated that about 70% of the isolates carried *ToxA* (Navathe et al., 2019). The presence of *ToxA* in Mexican *B. sorokiniana* isolates (Wu et al., 2021) indicates the similarity in virulence factors of pathogenic population in Mexico and SA that reflects why the selections in Agua Fria are effective for SA.

Three QTLs on 6A, 7A, and 7B in the CC population were identified only in SA. Anchoring the flanking markers of the 6A and 7B QTLs in the CS reference genome indicated that they both reside at the distal end of their respective chromosomes. Earlier, SB resistance QTL was mapped on 6A (Sharma et al., 2007a) and 7B (Singh et al., 2016; Ayana et al., 2018) but at different positions, implying that those mapped in the current study might be new. However, the QTL on 7A might not be new, since QTLs with similar confidence intervals were reported in the KATH × CIANO T79 population (Singh et al., 2018), in a CIMMYT wheat panel (Jamil et al., 2018), as well as in the Afghan landrace collection (Bainsla et al., 2020).

The pathotype diversity of *B. sorokiniana* and some climatic difference among SA and Mexico environments possibly play a significant role in the identification of effective QTLs. In our study, QTLs specific to SA were detected in both populations, providing an opportunity for breeding SB resistance cultivars in SA. In CC, an average of 27.2% reduction in SB was observed when the 5A and 5B QTLs were combined, compared to a 42.3% reduction of SB at Agua Fria (He et al., 2020). Detection of QTLs on 5A and *Sb2* at both SA and Agua Fria indicated the potentiality of these genes in resistance breeding for SA, but their lower phenotypic effects for SB in SA environment indicates the role of other QTLs, like those on 4D, 6A, and 7A. Furthermore, validation of these QTLs and markers over multiple locations and years will provide not only more insight into the SB resistance in SA but also more robust markers for the development of SB-resistant cultivars targeting SA.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: <https://hdl.handle.net/11529/10548547>.

AUTHOR CONTRIBUTIONS

PS, AJ, and XH designed the research and provided the plant material. VM and RC conducted the field experiments. CR, NG, and XH analyzed the data. CR drafted the first version of the manuscript. All authors contributed to the writing of the manuscript.

FUNDING

Financial support from the Bill and Melinda Gates Foundation and USAID through the CSISA, BGRI, DGGW projects, and

CGIAR Research Program on WHEAT for conducting this research is gratefully acknowledged. The financial support received by the first and second authors from the Indian Council of Agriculture Research (ICAR), India, is also acknowledged.

REFERENCES

- Ayana, G. T., Ali, S., Sidhu, J. S., Gonzalez-Hernandez, J. L., Turnipseed, B., and Sehgal, S. K. (2018). Genome-wide association study for spot blotch resistance in hard winter wheat. *Front. Plant Sci.* 9:926. doi: 10.3389/fpls.2018.00926
- Bainsla, N. K., Phuke, R. M., He, X., Gupta, V., Bishnoi, S. K., Sharma, R. K., et al. (2020). Genome-wide association study for spot blotch resistance in Afghan wheat germplasm. *Plant Pathol.* 69, 1161–1171. doi: 10.1111/ppa.13191
- Chand, R., Pandey, S., Singh, H., Kumar, S., and Joshi, A. K. (2003). Variability and its probable cause in natural populations of spot blotch pathogen *Bipolaris sorokiniana* of wheat (*T. aestivum* L.) in India / Variabilität und ihre wahrscheinliche Ursache in natürlichen Populationen des Erregers von Blattflecken (*Bipolaris sorokiniana*) an Weizen (*Triticum aestivum* L.) in Indien. *Z. Pflanzenkr. Pflanzenschutz J. Plant Dis. Prot.* 110, 27–35. doi: 10.14492/hokmj/1381758488
- Chaurasia, S., Chand, R., and Joshi, A. K. (2000). Relative dominance of *Alternaria tritici* Pras. et Prab. and *Bipolaris sorokiniana* (Sacc.) Shoemaker in different growth stages of wheat (*T. aestivum* L.). *J. Plant Dis. Prot.* 107, 176–181.
- Duveiller, E., Kandel, Y. R., Sharma, R. C., and Shrestha, S. M. (2005). Epidemiology of foliar blights (spot blotch and tan spot) of wheat in the plains bordering the Himalayas. *Phytopathology* 95, 248–256. doi: 10.1094/phyto-95-0248
- Friesen, T. L., Holmes, D. J., Bowden, R. L., and Faris, J. D. (2018). ToxA is present in the U.S. *Bipolaris sorokiniana* population and is a significant virulence factor on wheat harboring *Tsn1*. *Plant Dis.* 102, 2446–2452. doi: 10.1094/pdis-03-18-0521-re
- Gupta, P. K., Chand, R., Vasistha, N. K., Pandey, S. P., Kumar, U., Mishra, V. K., et al. (2018). Spot blotch disease of wheat: the current status of research on genetics and breeding. *Plant Pathol.* 67, 508–531. doi: 10.1111/ppa.12781
- Gurung, S., Mamidi, S., Bonman, J. M., Xiong, M., Brown-Guedira, G., and Adhikari, T. B. (2014). Genome-wide association study reveals novel quantitative trait loci associated with resistance to multiple leaf spot diseases of spring wheat. *PLoS One* 9:e108179. doi: 10.1371/journal.pone.0108179
- He, X., Dreisigacker, S., Sansaloni, C., Duveiller, E., Singh, R. P., and Singh, P. K. (2020). QTL mapping for spot blotch resistance in two bi-parental mapping populations of bread wheat. *Phytopathology* 110, 1980–1987. doi: 10.1094/phyto-05-20-0197-r
- Hetzler, J., Eyal, Z., Mehta, Y. R., Campos, L. A., Fehrmann, H., Kushnir, U., et al. (1991). "Interaction between spot blotch and wheat cultivars," in *Wheat for the Nontraditional Warm Areas: A Proceedings of the International Conference*, ed. D. A. Saunders (Foz do Iguaçu: CIMMYT), 125–145.
- Ibeagha, A. E., Hükelhoven, R., Schäfer, P., Singh, D. P., and Kögel, K. H. (2005). Model wheat genotypes as tools to uncover effective defense mechanisms against the hemibiotrophic fungus *Bipolaris sorokiniana*. *Phytopathology* 95, 528–532. doi: 10.1094/phyto-95-0528
- Jamil, M., Ali, A., Gul, A., Ghafoor, A., Ibrahim, A. M. H., and Mujeeb-Kazi, A. (2018). Genome-wide association studies for spot blotch (*Cochliobolus sativus*) resistance in bread wheat using genotyping-by-sequencing. *Phytopathology* 108, 1307–1314. doi: 10.1094/phyto-02-18-0047-r
- Joishi, A. K., Chand, R., and Arun, B. (2002). Relationship of plant height and days to maturity with resistance to spot blotch in wheat. *Euphytica* 123, 221–228.
- Joshi, A. K., and Chand, R. (2002). Variation and inheritance of leaf angle, and its association with spot blotch (*Bipolaris sorokiniana*) severity in wheat (*Triticum aestivum*). *Euphytica* 124, 283–291.
- Joshi, A. K., Ortiz-Ferrara, G., Crossa, J., Singh, G., Alvarado, G., Bhatta, M. R., et al. (2007a). Associations of environments in South Asia based on spot blotch disease of wheat caused by *Cochliobolus sativus*. *Crop Sci.* 47, 1071–1081. doi: 10.2135/cropsci2006.07.0477
- Joshi, A. K., Ortiz-Ferrara, G., Crossa, J., Singh, G., Sharma, R. C., Chand, R., et al. (2007b). Combining superior agronomic performance and terminal heat tolerance with resistance to spot blotch (*Bipolaris sorokiniana*) of wheat in the warm humid Gangetic plains of South Asia. *Field Crop Res.* 103, 53–61. doi: 10.1016/j.fcr.2007.04.010
- Kruijer, W., Boer, M. P., Malosetti, M., Flood, P. J., Engel, B., Kooke, R., et al. (2015). Marker-based estimation of heritability in immortal populations. *Genetics* 199, 379–398. doi: 10.1534/genetics.114.167916
- Kumar, S., Archak, S., Tyagi, R. K., Kumar, J., Vikas, V. K., Jacob, S. R., et al. (2016). Evaluation of 19,460 wheat accessions conserved in the Indian national genebank to identify new sources of resistance to rust and spot blotch diseases. *PLoS One* 11:e0167702. doi: 10.1371/journal.pone.0167702
- Kumar, S., Röder, M. S., Tripathi, S. B., Kumar, S., Chand, R., Joshi, A. K., et al. (2015). Mendelization and fine mapping of a bread wheat spot blotch disease resistance QTL. *Mol. Breed.* 35, 218.
- Kumar, U., Joshi, A. K., Kumar, S., Chand, R., and Röder, M. S. (2009). Mapping of resistance to spot blotch disease caused by *Bipolaris sorokiniana* in spring wheat. *Theor. Appl. Genet.* 118, 783–792. doi: 10.1007/s00122-008-0938-5
- Kumar, U., Joshi, A. K., Kumar, S., Chand, R., and Röder, M. S. (2010). Quantitative trait loci for resistance to spot blotch caused by *Bipolaris sorokiniana* in wheat (*T. aestivum* L.) lines 'Ning 8201' and 'Chirya 3'. *Mol. Breed.* 26, 477–491. doi: 10.1007/s11032-009-9388-2
- Kumar, U., Kumar, S., Prasad, R., Röder, M. S., Kumar, S., Chand, R., et al. (2019). Genetic gain on resistance to spot blotch of wheat by developing lines with near immunity. *Crop Breed. Genet. Genom.* 1:e190017. doi: 10.20900/cbagg20190017
- Li, H., Vikram, P., Singh, R. P., Kilian, A., Carling, J., Song, J., et al. (2015). A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. *BMC Genomics* 16:216. doi: 10.1186/s12864-015-1424-5
- Lillemo, M., Joshi, A. K., Prasad, R., Chand, R., and Singh, R. P. (2013). QTL for spot blotch resistance in bread wheat line Saar co-locate to the biotrophic disease resistance loci *Lr34* and *Lr46*. *Theor. Appl. Genet.* 126, 711–719. doi: 10.1007/s00122-012-2012-6
- Lu, P., Liang, Y., Li, D., Wang, Z., Li, W., Wang, G., et al. (2016). Fine genetic mapping of spot blotch resistance gene *Sb3* in wheat (*Triticum aestivum*). *Theor. Appl. Genet.* 129, 577–589. doi: 10.1007/s00122-015-2649-z
- Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL Ici mapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in bi-parental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001
- Navathe, S., Yadav, P. S., Chand, R., Mishra, V. K., Vasistha, N. K., Meher, P. K., et al. (2019). ToxA-Tsn1 interaction for spot blotch susceptibility in Indian wheat: an example of inverse gene-for-gene relationship. *Plant Dis.* 104, 71–81. doi: 10.1094/pdis-05-19-1066-re
- Rosyara, U. R., Subedi, S., Duveiller, E., and Sharma, R. C. (2010). The effect of spot blotch and heat stress on variation of canopy temperature depression, chlorophyll fluorescence and chlorophyll content of hexaploid wheat genotypes. *Euphytica* 174, 377–390. doi: 10.1007/s10681-010-0136-9
- Saari, E. E., and Prescott, J. M. (1975). A scale for appraising the foliar intensity of wheat disease. *Plant Dis. Rep.* 59, 377–380.
- Sharma, R. C., Duveiller, E., Ahmed, F., Arun, B., Bhandari, D., Bhatta, M. R., et al. (2004). Helminthosporium leaf blight resistance and agronomic performance of wheat genotypes across warm regions of South Asia. *Plant Breed.* 123, 520–524. doi: 10.1111/j.1439-0523.2004.01006.x
- Sharma, R. C., Duveiller, E., and Jacquemin, J. M. (2007a). Microsatellite markers associated with spot blotch resistance in spring wheat. *J. Phytopathol.* 155, 316–319. doi: 10.1111/j.1439-0434.2007.01238.x
- Sharma, R. C., Duveiller, E., and Ortiz-Ferrera, G. (2007b). Progress and challenge towards reducing wheat spot blotch threat in the Eastern Gangetic Plains of South Asia: is climate change already taking its toll? *Field Crop Res.* 103, 109–118. doi: 10.1016/j.fcr.2007.05.004

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.641324/full#supplementary-material>

- Siddique, A. B., Hossain, M. H., Duveiller, E., and Sharma, R. C. (2006). Progress in wheat resistance to spot blotch in Bangladesh. *J. Phytopathol.* 154, 16–22. doi: 10.1111/j.1439-0434.2005.01049.x
- Singh, P. K., He, X., Sansaloni, C. P., Juliana, P., Dreisigacker, S., Duveiller, E., et al. (2018). Resistance to spot blotch in two mapping populations of common wheat is controlled by multiple QTL of minor effects. *Int. J. Mol. Sci.* 19:4054. doi: 10.3390/ijms19124054
- Singh, P. K., Zhang, Y., He, X., Singh, R. P., Chand, R., Mishra, V. K., et al. (2015). Development and characterization of the 4th CSISA-spot blotch nursery of bread wheat. *Eur. J. Plant Pathol.* 143, 595–605. doi: 10.1007/s10658-015-0712-x
- Singh, V., Singh, G., Chaudhury, A., Ojha, A., Tyagi, B. S., Chowdhary, A. K., et al. (2016). Phenotyping at hot spots and tagging of QTLs conferring spot blotch resistance in bread wheat. *Mol. Biol. Rep.* 43, 1293–1303. doi: 10.1007/s11033-016-4066-z
- Veldboom, L. R., and Lee, M. (1996). Genetic mapping of quantitative trait loci in maize in stress and non-stress environments: I. Grain yield and yield components. *Crop Sci.* 36, 1310–1319. doi: 10.2135/cropsci1996.0011183x003600050040x
- 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
- Wu, L., He, X., Lozano, N., Zhang, X., and Singh, P. K. (2021). *ToxA*, a significant virulence factor involved in wheat spot blotch disease, exists in the Mexican population of *Bipolaris sorokiniana*. *Trop. Plant Pathol.* doi: 10.1007/s40858-020-00391-4
- Zhang, P., Guo, G., Wu, Q., Chen, Y., Xie, J., Lu, P., et al. (2020). Identification and fine mapping of spot blotch (*Bipolaris sorokiniana*) resistance gene *Sb4* in wheat. *Theor. Appl. Genet.* 133, 2451–2459. doi: 10.1007/s00122-020-03610-3
- Zhu, Z., Bonnett, D., Ellis, M., Singh, P., Heslot, N., Dreisigacker, S., et al. (2014). Mapping resistance to spot blotch in a CIMMYT synthetic-derived bread wheat. *Mol. Breed.* 34, 1215–1228. doi: 10.1007/s11032-014-0111-6

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 Roy, Gahtyari, He, Mishra, Chand, Joshi 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.



Genome-Wide Association Studies Reveal All-Stage Rust Resistance Loci in Elite Durum Wheat Genotypes

Meriem Aoun^{1*}, Matthew N. Rouse^{2,3}, James A. Kolmer^{2,3}, Ajay Kumar¹ and Elias M. Elias^{1*}

¹ Department of Plant Sciences, North Dakota State University, Fargo, ND, United States, ² Cereal Disease Laboratory, United States Department of Agriculture–Agricultural Research Service, St. Paul, MN, United States, ³ Department of Plant Pathology, University of Minnesota, St. Paul, MN, United States

OPEN ACCESS

Edited by:

Anna Maria Mastrangelo,
Research Centre for Cereal
and Industrial Crops, Council
for Agricultural Research
and Economics (CREA), Italy

Reviewed by:

Marco Maccaferri,
University of Bologna, Italy
Fernando Martinez,
University of Seville, Spain

*Correspondence:

Meriem Aoun
ma792@cornell.edu
Elias M. Elias
elias.elias@ndsu.edu

Specialty section:

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

Received: 12 December 2020

Accepted: 17 March 2021

Published: 12 April 2021

Citation:

Aoun M, Rouse MN, Kolmer JA,
Kumar A and Elias EM (2021)
Genome-Wide Association Studies
Reveal All-Stage Rust Resistance Loci
in Elite Durum Wheat Genotypes.
Front. Plant Sci. 12:640739.
doi: 10.3389/fpls.2021.640739

Leaf rust, caused by *Puccinia triticina* (Pt), stripe rust caused by *Puccinia striiformis* f. sp. *tritici* (Pst), and stem rust caused by *Puccinia graminis* f. sp. *tritici* (Pgt) are major diseases to wheat production globally. Host resistance is the most suitable approach to manage these fungal pathogens. We investigated the phenotypic and genotypic structure of resistance to leaf rust, stem rust, and stripe rust pathogen races at the seedling stage in a collection of advanced durum wheat breeding lines and cultivars adapted to Upper Mid-West region of the United States. Phenotypic evaluation showed that the majority of the durum wheat genotypes were susceptible to Pt isolates adapted to durum wheat, whereas all the genotypes were resistant to common wheat type-Pt isolate. The majority of genotypes were resistant to stripe rust and stem rust pathogen races. The durum panel genotyped using Illumina iSelect 90 K wheat SNP assay was used for genome-wide association mapping (GWAS). The GWAS revealed 64 marker-trait associations (MTAs) representing six leaf rust resistance loci located on chromosome arms 2AS, 2AL, 5BS, 6AL, and 6BL. Two of these loci were identified at the positions of *Lr52* and *Lr64* genes, whereas the remaining loci are most likely novel. A total of 46 MTAs corresponding to four loci located on chromosome arms 1BS, 5BL, and 7BL were associated with stripe rust response. None of these loci correspond to designated stripe rust resistance genes. For stem rust, a total of 260 MTAs, representing 22 loci were identified on chromosome arms 1BL, 2BL, 3AL, 3BL, 4AL, 5AL, 5BL, 6AS, 6AL, 6BL, and 7BL. Four of these loci were located at the positions of known genes/alleles (*Sr7b*, *Sr8155B1*, *Sr13a*, and *Sr13b*). The discovery of known and novel rust resistance genes and their linked SNPs will help diversify rust resistance in durum wheat.

Keywords: leaf rust, stripe rust, stem rust, durum wheat, resistance, association mapping, molecular markers

INTRODUCTION

Durum wheat [$2n = 4x = 28$, AABB, *Triticum turgidum* L. Var. *durum* (Desf.)] is the second most cultivated wheat crop. It accounts for about 8% of the world's total wheat production (Mengistu and Pè, 2016) and is mainly produced in the Mediterranean region, Eastern Europe, and North America (Royo et al., 2009). Annual worldwide durum wheat production is estimated to be around

36 million tons (Magallanes-López et al., 2017), with approximately, 2.5 million tons produced in the United States. North Dakota's production of durum wheat accounts for over 50% of total U.S. production (NASS, 2018). Leaf rust, stripe rust and stem are major fungal diseases threatening durum wheat production globally. Rust resistance is the most environmentally and economically feasible approach for managing these diseases. Therefore, the development and deployment of rust resistant cultivars is a major goal of wheat breeding programs worldwide.

Tetraploid durum wheat has historically been more resistant to leaf rust than hexaploid common wheat (*T. aestivum* L.) (Singh et al., 2004; Herrera-Foessel et al., 2006). However, during the last 20 years, new durum-type *Pt* races have emerged and caused leaf rust epidemics in several durum wheat producing regions (Singh et al., 2004; Goyeau et al., 2012; Mishra et al., 2015; Aoun et al., 2016). Virulent *Pt* isolates have not been found yet in North Dakota, however, a highly virulent race (BBBQJ) was reported in California and Kansas, United States (Kolmer, 2013, 2015a). This poses a threat to the major durum-producing regions of the USA and Canada. *Pt*-isolates from durum wheat are often avirulent to most leaf rust resistance (*Lr*) genes in common wheat (Huerta-Espino and Roelfs, 1992; Ordoñez and Kolmer, 2007a). The durum wheat type-*Pt* isolates from several durum wheat producing countries have similar phenotypic reactions on 'Thatcher' wheat near-isogenic lines and similar or identical SSR and SNP genotypes, suggesting a common origin (Ordoñez and Kolmer, 2007a,b; Aoun et al., 2020; Kolmer et al., 2020). Other *Pt*-isolates collected on tetraploid wheat in Ethiopia (designated as race EEEEE) are avirulent to Thatcher wheat and have a unique molecular genotype compared to all other *Pt*-isolates collected from durum wheat and common wheat globally (Ordoñez and Kolmer, 2007a,b; Kolmer and Acevedo, 2016; Aoun et al., 2020; Kolmer et al., 2020).

A total of 79 *Lr* genes have been identified in wheat, only 20 of them are known to be present in durum wheat (Desiderio et al., 2014; Qureshi et al., 2018). In response to leaf rust epidemics in many durum producing countries, a number of *Lr* genes were identified in this crop including *Lr3a*, *Lr14a*, *Lr27+Lr31*, *Lr61*, *Lr79*, and *LrCamayo* (Herrera-Foessel et al., 2007, 2008a,b; Huerta-Espino et al., 2009; Qureshi et al., 2018). Other not yet cataloged *Lr* genes were also detected in durum wheat landraces and cultivars (Loladze et al., 2014; Aoun et al., 2016, 2017, 2019; Kthiri et al., 2018, 2019). However, due to continuous virulence evolution of *Pt* isolates on many of the identified *Lr* genes, diversifying the genetic basis for leaf rust resistance in durum wheat is a priority.

Stripe rust is another major disease of wheat worldwide (Chen, 2005). Aggressive *Pst* races adapted to high temperatures have emerged and spread into most wheat producing regions (Milus et al., 2009). Over 80 stripe rust resistance (*Yr*) genes have been designated in wheat (McIntosh et al., 2013, 2017, Gessese et al., 2019). The *Yr* genes that were derived from tetraploid wheat (*T. turgidum* L. ssp) include *Yr7*, *Yr15*, *Yr24/Yr26*, *Yr30*, *Yr35*, *Yr36*, *YrH52*, *Yr53*, *Yr64*, and *Yr65* (McFadden, 1939; Macer, 1966; McIntosh and Lagudah, 2000;

Peng et al., 2000; Ma et al., 2001; Marais et al., 2005; Uauy et al., 2005; Xu et al., 2013; Cheng et al., 2014). However, most of the *Yr* genes identified in wheat are race specific and have become ineffective against the rapidly evolving pathogen (Chen, 2013; McIntosh et al., 2013; Rosewarne et al., 2013). Therefore, identification and pyramiding of new genes is needed for more effective management of this rapidly evolving pathogen.

Stem rust has historically threatened common wheat and durum wheat production. The Ug99 race group that appeared in East Africa overcame several widely used wheat stem rust resistance (*Sr*) genes (Jin et al., 2007; Singh et al., 2011). More than 70 cataloged *Sr* genes have been characterized in durum and common wheat. Only 31 genes are still effective against at least one race of the 13 Ug99 variants (Rouse et al., 2011, 2014; Singh et al., 2011, 2015). Approximately half of these genes were introgressed into wheat from secondary and tertiary gene pools (Rouse et al., 2014; Singh et al., 2015) and only a few genes have been identified in durum wheat. Designated *Sr* genes that have been reported in tetraploid wheat include *Sr2*, *Sr7a*, *Sr8b*, *Sr8155B1*, *Sr9d*, *Sr9e*, *Sr9g*, *Sr11*, *Sr12*, *Sr13a*, *Sr13b*, *Sr14*, and *Sr17* (Jin et al., 2007; Singh et al., 2015; Nirmala et al., 2017; Saini et al., 2018; Zhang et al., 2017).

In North American durum wheat cultivars, resistance to the Ug99 lineage is mainly due to *Sr13*, of which the *Sr13a* allele was first identified in Khapstein, a hexaploid wheat derivative of cultivated emmer wheat (*T. turgidum* L. ssp. *dicoccum*, $2n = 4x = 28$, AABB) cv. Khapli (Knott, 1962; Jin et al., 2007; Klindworth et al., 2007; Zhang et al., 2017). *Sr9e* is also another *Sr* gene frequently deployed in durum wheat (Olivera et al., 2012; Saini et al., 2018). Nirmala et al. (2017) recently identified a possible *Sr8* allele, designated as *Sr8155B1*, in the durum wheat line '8155-B1.' *Sr8155B1* is effective to a variant of the Ug99 race TTKST but not to race TTKSK (Nirmala et al., 2017). However, the frequency of this allele in durum wheat cultivars is not yet determined. Besides the Ug99 race group, additional *Pgt*-races with broad virulence spectra have also emerged during the last decade including TRTTF, JRCQC, and TKTTF. These races do not belong to the Ug99 lineage and pose serious threat to common wheat and durum production (Olivera et al., 2012, 2015). Among these races, TRTTF and JRCQC were reported to be virulent to the major known components of stem rust resistance in North American durum cultivars *Sr13* and *Sr9e* (Olivera et al., 2012). However, according to Zhang et al. (2017), *Sr13a* is effective to both JRCQC and TRTTF, and *Sr13b* is effective to TRTTF, but not JRCQC. Identifying and characterizing new sources of stem rust resistance in durum wheat is needed to manage future outbreaks.

This study was designed: (1) to determine levels of leaf rust, stem rust, and stripe rust resistance in a large collection of elite durum wheat lines at seedling stage, (2) to determine the genetic architecture of rust resistance loci using GWAS and Infinium 90K wheat SNP assay (3) to detect novel seedling resistance (all-stage resistance) loci to *Pt*, *Pst*, and *Pgt* races that could be used in breeding programs, and (4) to identify SNPs associated with seedling rust resistance loci for marker assisted breeding.

MATERIALS AND METHODS

Plant Materials

A collection of 248 durum wheat genotypes was used in this study. The collection represented advanced breeding lines evaluated in the North Dakota State University's (NDSU) Uniform Regional Durum Nursery (URDN) from 1997 to 2014 (for more details, see Johnson et al., 2019; **Supplementary Table 1**). These genotypes were regularly evaluated for agronomic and quality traits over the years in multiple environments. Thus, this plant material represents the core of the NDSU's durum breeding program.

Leaf Rust Phenotyping

The durum wheat collection was screened at the seedling stage with six *Pt* isolates (**Supplementary Table 1**). Five of these isolates (TUN 20-1, ETH 13D17-1, MEX10, ETH 63-1, and MOR 33-1) were durum wheat type isolates, while ALK-ND is a common wheat type isolate from North Dakota. The virulence/avirulence phenotypes of the *Pt* isolates were based on the infection types (ITs) of 20 Thatcher near-isogenic lines (NILs) at seedling stage as described by Long and Kolmer (1989). The Tunisian (TUN 20-1) and Moroccan (MOR 33-1) isolates were both of race BBBSJ (virulent to the *Lr* genes *LrB*, *Lr10*, *Lr14a*, *Lr14b*, and *Lr20*). The Mexican isolate MEX10 was of race BBBQJ (virulent to the *Lr* genes *LrB*, *Lr10*, *Lr14b*, and *Lr20*). The two Ethiopian isolates ETH 63-1 and ETH 13D17-1 designated as race EEEEE are avirulent on the Thatcher wheat. The common wheat type isolate ALK-ND, designated as race MBDSS was isolated from the durum wheat cultivar 'Alkabo' (PI 642020) in North Dakota and is virulent to the *Lr* genes *Lr1*, *Lr3a*, *Lr3bg*, *Lr10*, *Lr14a*, *Lr14b*, *Lr17*, *Lr20*, and *LrB*.

The phenotyping using isolates EEEEE_ETH 63-1, BBBSJ_MOR 33-1, and MBDSS_ALK-ND was performed at the biosafety level-2 facility at the Agricultural Experiment Station Greenhouse Complex, Fargo, ND, United States using a randomized complete block design (RCBD) with two replicates. In each replicate five-to-seven plants/line were tested and the common wheat cultivar Thatcher and the leaf rust susceptible durum line 'RL6089' were included twice as checks in each of the 50-cell trays. For each experiment, two replicates of Thatcher NIL differentials were included to confirm the virulence phenotype of *Pt*-isolates. Seedling growth conditions, inoculum increase and preparation, inoculation, and greenhouse conditions under which the inoculated plants were kept until disease phenotyping were as described by Aoun et al. (2019).

The screening experiments with the remaining three isolates EEEEE_ETH 13D17-1, BBBQJ_MEX10, and BBBSJ_TUN 20-1 were done at the U.S. Department of Agriculture- Agricultural Research Service (USDA-ARS), Cereal Disease Laboratory (CDL) in Saint Paul, MN, United States. The seedling tests using these three isolates were performed in a single replicate with five-to-seven plants/line. The common wheat Thatcher and the durum line RL6089 were included as checks. The detailed protocols of plant growing conditions and inoculation were described in Kolmer and Hughes (2013).

Leaf rust ITs were taken 12–14 days after inoculation on the second leaf using a 0–4 scale (Long and Kolmer, 1989; McIntosh et al., 1995) where IT of '0' = no visible symptoms, IT of ';' = hypersensitive flecks, IT of '1' = small uredinia with necrosis, IT of '2' = small-to medium-size uredinia surrounded by chlorosis, IT of '3' = medium-size uredinia with no chlorosis or necrosis, and IT of '4' = large uredinia with no necrosis or chlorosis. Larger and smaller uredinia than expected for each IT were designated with '+' and '-', respectively. Seedling plants exhibiting ITs of 0–2⁺ and 'X' (a mixture of resistant and susceptible reactions evenly distributed on the leaf surface) were considered resistant, whereas seedling plants with ITs of 3–4 were considered susceptible (Long and Kolmer, 1989; McIntosh et al., 1995). In situations where multiple ITs were observed on the same leaf surface, the plant reaction was recorded as the most predominant IT followed by the least predominant IT.

Stripe Rust Phenotyping

Three *Pst* races (PSTv-37, PSTv-41, and PSTv-52) collected from common wheat in North Dakota (**Supplementary Table 2**) were used to screen the durum genotypes. These three *Pst* races are the only ones currently present in North Dakota. PSTv-37 has been the most widely distributed race across the United States (Wan et al., 2016) and has a virulence/avirulence phenotype of Yr6, 7, 8, 9, 17, 27, 43, 44, Tr1, Exp2/Yr1, 5, 10, 15, 24, 32, SP, 76. The race PSTv-52 that has been widely distributed in the United States¹ has a virulence/avirulence profile of Yr6, 7, 8, 9, 17, 27, 43, 44, Exp2/Yr1, 5, 10, 15, 24, 32, SP, Tr1, 76. The race PSTv-41 is considered the most virulent race in ND and has a virulence/avirulence profile of Yr6, 7, 8, 9, 10, 17, 24, 27, 32, 43, 44, Tr1, Exp2/Yr1, 5, 15, SP, 76.

To screen for stripe rust, three separate experiments (one experiment/*Pst* race) with the same set of durum genotypes ($n = 248$) were planted at the Fargo Agricultural Experiment Station Greenhouse Complex, ND, United States. In each experiment, five-to-seven seeds/genotype were planted in 50-well trays. The susceptible cultivar 'Avocet' was included twice in each tray as check. To confirm the race identity, a set of 18 differential lines containing each a single *Yr* gene (Wan and Chen, 2014) was included alongside each single-race experiment. The seedlings were grown in a rust-free greenhouse at 22°C/18°C (day/night) and 16 h photoperiod. When the second leaves were fully expanded, the plants were spray inoculated with fresh rust urediniospores suspended in Soltrol-170 oil (Phillips Petroleum, Bartlesville, OK, United States) at a concentration of 0.01 g/mL. After the Soltrol-170 oil dried on the leaf surface, the inoculated plants were incubated in a dark dew chamber at 10°C with 100% relative humidity for 24 h. The seedlings were later transferred to a rust-free incubated growth chamber at 17°C/8°C (day/night) and 16 h photoperiod. The seedling ITs were recoded 16–18 days post-inoculation on a scale of 0–9 (Line and Qayoum, 1992). IT of '0' = no visible signs or symptoms, IT of '1' = necrotic or chlorotic flecks with no sporulation; IT of '2' = necrotic and/or chlorotic blotches or stripes with no sporulation; IT of '3' = necrotic and/or chlorotic blotches or stripes with only a

¹<http://striperust.wsu.edu>

trace of sporulation; IT of '4', '5', and '6' corresponds to necrotic and/or chlorotic blotches or stripes with light, intermediate, and moderate sporulation, respectively; and IT of '7', '8', and '9' corresponds to abundant sporulation with necrotic and/or chlorotic stripes or blotches, chlorosis around the sporulation area, and no chlorosis or necrosis, respectively. ITs from 0 to 3 were considered resistant responses, ITs from 4 to 6 were considered intermediate responses and ITs from 7 to 9 were considered susceptible responses.

Stem Rust Phenotyping

The durum wheat genotypes were tested at seedling stage with six African *Pgt* races TTKSK (isolate 04KEN156/04; Jin and Singh, 2006), TTKST (06KEN19v3; Jin et al., 2008), TTKTT (14KEN58-1; Newcomb et al., 2016), TKTTF (13ETH18-1; Olivera et al., 2015), TRTTF (06YEM34-1; Olivera et al., 2012), and JRCQC (09ETH08-3; Olivera et al., 2012) (**Supplementary Table 3**). The durum lines were phenotyped in the biosafety level-3 facility at the USDA-ARS CDL in St. Paul, MN, United States. The lines were planted in two replicates corresponding to different experiments with different planting and inoculation dates. Five seedlings per line were planted per replicate for all six *Pgt* races. The inoculum preparation, inoculation, greenhouse conditions, and disease screening were as described by Hundie et al. (2019). In brief, the urediniospores stored at -80°C were heat shocked at 45°C for 15 min. For inoculation, gelatin capsules including 14 mg spores were suspended in 0.75 ml mineral oil (Sotrol 170, Phillips Petroleum, Borger, TX, United States) and sprayed onto the plant primary leaves of 240 wheat seedling plants corresponding to 48 wheat lines. After the Sotrol-170 oil dried on the leaf surface, the inoculated plants were placed in a humidity chamber in the dark for 14-to-16 h at 22°C , then exposed to high pressure sodium vapor lamps for 3–4 h. The plants were then transferred to the greenhouse and kept at temperature of 19 – 22°C and 16 h photoperiod for 10–12 days. The seedling ITs were scored using the Stakman 0–4 scale (Stakman et al., 1962). Plants with ITs of 0–2+3 were considered resistant and those with IT of 3–4 were considered susceptible.

Phenotypic Data Analysis

For statistical analysis, the 0–4 scale for leaf rust and stem rust screening was converted to a linearized 0–9 scale (Zhang et al., 2014) where plants with ITs of 0–6 were classified as resistant and those with ITs of 7–9 were considered susceptible. For further analysis, the mean of replicates per trait were used. Pairwise Pearson's correlations between traits were calculated and plotted using the 'corrplot' package (Wei and Simko, 2013) in the software R 3.4.1 (R Core Team, 2016). Correlation values were considered significantly different from zero at P -value ≤ 0.05 .

Genotyping

The durum collection was genotyped as described by Johnson et al. (2019) using the Illumina iSelect 90K wheat SNP assay (Wang et al., 2014). The 90K wheat SNP assay generated 17,377 polymorphic SNPs. Markers which were in common with those included in the tetraploid wheat consensus map (Maccaferri et al., 2015) were kept for further analysis (**Supplementary Table 4**).

Additionally, a diagnostic marker for the presence of either *Sr13* allele (Zhang et al., 2017), a linked marker to *Sr8155B1* (Nirmala et al., 2017), and three dCAPS markers used to discriminate *Sr13a* and *Sr13b* were also used to genotype the durum wheat collection. The durum wheat collection was genotyped using derived cleaved amplified polymorphic sequence (dCAPS) markers for *Sr13* and its three alleles R1 (*Sr13a-R1*), R2 (*Sr13b*), and R3 (*Sr13a-R3*). Markers *dCAPS_Sr13* (Zhang et al., 2017), *dCAPS_Sr13_R1cut*, *dCAPS_Sr13_R2nocut*, and *dCAPS_Sr13_R3nocut* were used to identify *Sr13*, *Sr13a-R1*, *Sr13b*, *Sr13a-R3*, respectively (**Supplementary Table 5**). *Sr13a-R1* and *Sr13a-R3* correspond to the two resistant haplotypes of *Sr13a*: R1 and R3 (Zhang et al., 2017). The dCAPs markers used to discriminate among the two *Sr13* alleles were designed based on the sequence information of the resistant haplotypes of *Sr13* in Zhang et al. (2017). The primer sequences of *Sr13* gene/alleles, the restriction enzymes (RE), and the resulting PCR product sizes after RE digestion are described in **Supplementary Table 5**. The Kompetitive Allele Specific PCR (KASP) marker (*KASP_6AS_IWB10558*) was used to postulate the presence of the gene *Sr8155B1* (Nirmala et al., 2017). Heterozygotes were converted into missing data. Polymorphic markers with $>10\%$ missing data and minor allele frequency (MAF) $< 3\%$ were excluded from further analysis.

Linkage Disequilibrium and Population Structure

Linkage disequilibrium (LD) was performed using JMP Genomics 8.1 software (SAS Institute Inc, 2004) as described by Aoun et al. (2016). The LD estimates for intrachromosomal markers were calculated as the squared correlation coefficient (R^2) for each of the marker pairs. The genome-wide LD decay was estimated by plotting LD estimates (R^2) from all 14 durum wheat chromosomes against the corresponding pairwise genetic distances in cM. The genetic positions of the markers were according to the durum wheat consensus map of Maccaferri et al. (2015). Smoothing spline fit was applied to LD decay plot.

The principal component analysis (PCA) was used to examine the population structure (Q matrix). SNPs with LD (R^2) ≤ 0.2 were used to estimate the Q matrix. The identity-by-state (IBS) matrix or Kinship matrix (K matrix) that represents the proportion of shared alleles for all pairwise comparisons between genotypes was also estimated. The K and Q matrices were estimated using JMP Genomics 8.1 software.

Genome-Wide Association Analysis

For each trait, mixed linear model for genome-wide association analyses were performed using JMP Genomics 8.1 software. Five regression models were tested to identify the best model per trait from which MTAs will be derived. The tested models include (i) naïve, (ii) kinship, (iii) kinship plus population structure (first two PCs), (iv) kinship plus population structure (first three PCs), and (v) kinship plus population structure (first four PCs). The K and the Q matrices were included in the genome-wide association analysis model to reduce the chance of false-positive MTAs. Each of the markers was fitted into the regression equation

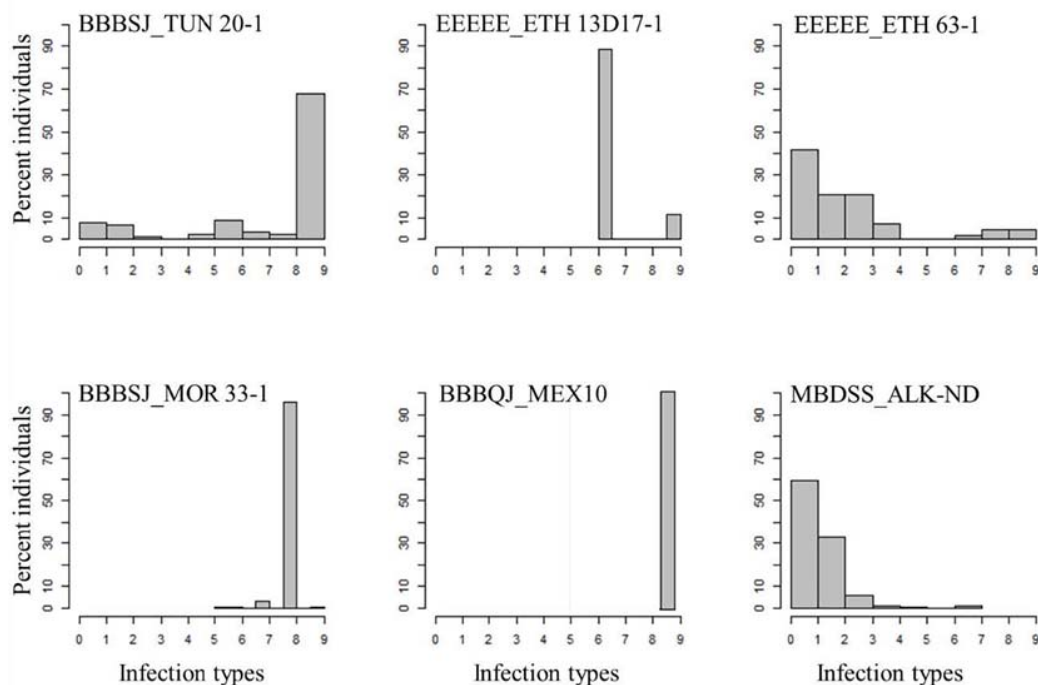


FIGURE 1 | Distributions of the seedling responses of the durum wheat genotypes to *Puccinia triticina* isolates BBBSJ_TUN 20-1, EEEEE_ETH 13D17-1, EEEEE_ETH 63-1, BBBSJ_MOR 33-1, BBBQJ_MEX10, and MBDSS_ALK-ND. X-axis corresponds to linearized Stakman scale (0-to-9).

to generate a *P*-value. The best association mapping model (of the five tested regression models) was selected based on the Bayesian Information Criterion (BIC), where the lowest BIC value corresponded to the best model (Ghosh et al., 2006; Zhang et al., 2010). For each trait, the marker *P*-values of the selected model were used to calculate the *P*-value of the false positive discovery rate (FDR) (Benjamini and Yekutieli, 2001). MTAs were considered significant at *P*-value of $FDR \leq 0.05$. The LD estimates between significant markers and marker genetic positions on the tetraploid consensus map (Maccaferri et al., 2015) were used to group MTAs from the GWAS into the same or different underlying loci. Each locus was represented by the most significant SNP marker. The physical and genetic position of the most significant marker per locus and any markers from the literature used for comparative mapping was based on the durum wheat cv. Svevo genome v1 (Maccaferri et al., 2019) and the tetraploid consensus map (Maccaferri et al., 2015), respectively. In the case of multiple identified loci on the same chromosome, the loci were ordered according to their most significant SNP genetic positions on the tetraploid consensus map of Maccaferri et al. (2015).

RESULTS

Phenotypic Data

Leaf Rust

All the durum wheat genotypes were resistant to the common wheat type isolate MBDSS_ALK-ND. For the *Pt* durum wheat

type isolates, the percentage of susceptible lines varied depending on the isolate (**Supplementary Table 1**). For instance, 10% of the genotypes were susceptible to the Ethiopian isolate EEEEE_ETH 63-1, while 28% of the genotypes showed susceptibility to the Ethiopian isolate EEEEE_13D17-1. The distribution of the ITs to EEEEE_13D17-1 was bimodal, where two ITs were observed. A total of 72% of the genotypes exhibited a mesothetic reaction (IT = '3⁺'), while the remaining genotypes showed IT = '3⁺'. The plant reactions to EEEEE_ETH 63-1 ranged between '3' and 3⁺. Even though the two Ethiopian isolates had similar race designation EEEEE (avirulent to the common wheat cv. Thatcher), they carried different virulence/avirulence phenotypes to the durum genotypes in our study (**Figure 1** and **Supplementary Table 1**).

In contrast to the Ethiopian isolates, the percentages of susceptible genotypes to durum wheat type isolates from Morocco, Tunisia, and Mexico were much higher. For instance, all the durum genotypes were susceptible to isolate BBBQJ_MEX10. Similarly, 74 and 98% of the genotypes were susceptible to isolates BBBSJ_TUN 20-1 and BBBSJ_MOR 33-1, respectively. The most resistant lines to race BBBSJ_MOR 33-1 had IT of '23', whereas the most resistant lines to race BBBSJ_TUN 20-1 had IT of '3', suggesting that these two isolates of the same race (based on Thatcher wheat differentials) carried different virulence/avirulence profiles to durum wheat (**Figure 1** and **Supplementary Table 1**).

The top four durum wheat cultivars grown in ND in 2019 were Joppa (PI 673106, 30.2%), Divide (PI 642021, 21.2%), Alkabo (PI 642020, 7.8%), and Carpio (PI 670039, 6.1%)

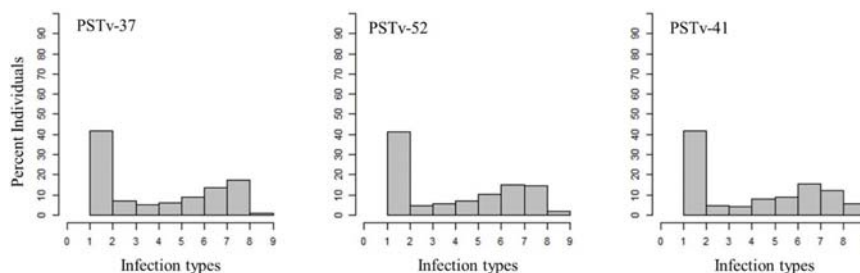


FIGURE 2 | Distributions of the seedling responses of the durum wheat genotypes to *Puccinia striiformis* f. sp. *tritici* races PSTv-37, PSTv-41, and PSTv-52.

(USDA, NASS, North Dakota Field Office, 2019). All of these cultivars were resistant to EEEEE_ETH 13D17-1, EEEEE_ETH 63-1, and MBDSS_ALK-ND but susceptible to the Mexican isolate BBBQJ_MEX10. Joppa showed intermediate IT to BBBSJ_TUN 20-1 (IT = '23') and to BBBSJ_MOR 33-1 (IT = '32⁺'). Divide was resistant to BBBSJ_TUN 20-1, whereas Alkabo and Carpio were susceptible to this Tunisian isolate. Divide was resistant to BBBSJ_MOR 33-1, while Alkabo and Carpio were susceptible to this isolate (**Supplementary Table 1**).

Stripe Rust

A total of 69% of the durum wheat genotypes were resistant to races PSTv-37 and PSTv-52, while 67% of the lines were resistant to race PSTv-41. The ITs to the three *Pst* races ranged between 1 and 9. The cultivars Divide, Alkabo, Carpio were resistant to all the three *Pst* races. Joppa was resistant to races PSTv-37 and PSTv-41 but not to PSTv-52 (**Figure 2** and **Supplementary Table 2**).

Stem Rust

About 81–99% of the genotypes were resistant to the three Ug99-lineage races TTKSK, TTKST, and TTKTT. For race TTKSK, the ITs ranged from 1 to 3⁺ with most of the lines showing IT of '2⁺'. The ITs to races TTKST and TTKTT ranged between 0; and 33⁺ with most of the genotypes showing IT = '0;'. For race TKTTE, only the breeding line 'D07726' showed a susceptible IT, while the remaining genotypes showed resistant ITs that ranged between '0;' and '2' with the most frequent resistant IT = '0;' (**Figure 3** and **Supplementary Table 3**).

All the durum wheat genotypes showed resistant ITs to race TRTTF ranging between '0;' and '2.' Like races TTKST, TTKTT, and TKTTE, the most frequent resistant IT to race TRTTF was '0;'. Even though, 99–100% of the durum genotypes were resistant to race TTKST, TTKTT, TKTTE, and TRTTF, there were phenotypic variations within the resistant ITs (**Figure 3** and **Supplementary Table 3**) appropriate to conduct further analysis (e.g., GWAS). Of all the *Pgt* races used for screening, race JRCQC was the most virulent race on the durum wheat collection, with 44% of the genotypes showing susceptibility. The resistant ITs to JRCQC ranged from '1;' to '2⁺3' with most of the resistant genotypes showing ITs of '22⁺' to '2⁺3' (**Figure 3** and **Supplementary Table 3**). The durum cultivars Carpio and Alkabo showed resistance to all *Pgt*-races. Divide was resistant to

all races except TTKSK and JRCQC, while Joppa was resistant to all races except TTKSK (**Supplementary Table 3**).

Phenotypic Data Correlations

For correlation analyses, we considered only traits with phenotypic variations (**Figure 4**). Pearson's correlation between linearized ITs showed a significant correlation ($r = 0.8$, $P\text{-value} \leq 0.05$) between the durum genotype responses to the Ethiopian *Pt* races EEEEE_ETH 63-1 and EEEEE_13D17-1. However, there were no significant correlations between the ITs to BBBSJ_TUN 20-1 and the ITs to both Ethiopian isolates of race EEEEE. There were strong significant correlations between ITs to the three *Pst* races that ranged between 0.8 and 0.9. For *Pgt* races, we observed significant correlations ($r = 0.7\text{--}0.9$, $P\text{-value} \leq 0.05$) between ITs to races TTKST, TTKTT, TKTTE, and TRTTF. ITs to TTKSK and JRCQC were not significantly correlated with ITs to any of the remaining four *Pgt*-races. There was no correlation between ITs to TTKSK and JRCQC. We found no significant correlations between ITs to different rust pathogens, suggesting that different genetic loci confer resistance to leaf rust, stripe rust, and stem rust in this durum wheat collection (**Figure 4**).

Marker Properties and Linkage Disequilibrium Analysis

After marker filtering, 10,891 SNPs included in the tetraploid wheat consensus map with MAF $\geq 3\%$ and missing data points $\leq 10\%$ were used for further analysis. Of the 10,891 SNPs, there were 4,779 (43.9%) SNPs on the genome A and 6,112 (56.1%) SNPs on the genome B. Additional four diagnostic dCAPS markers for *Sr13* gene/alleles and a single KASP marker for *Sr8155B1* gene were included. The genome-wide linkage disequilibrium (LD) dropped by half to 0.33 within 2.5 cM on average (**Figure 5**). Therefore, MTAs from the GWAS within 2.5 cM on average and with LD (R^2) ≥ 0.3 were considered underlying the same locus. In addition, we considered the pairwise LD (R^2 cutoff = 0.3) between significant markers on the same chromosome arm to identify the loci.

Kinship Analysis, Population Structure, and Regression Model Selection for GWAS

For the identity-by-state matrix or kinship matrix (*K* matrix), there were some hotspots (red color in the heat map) between

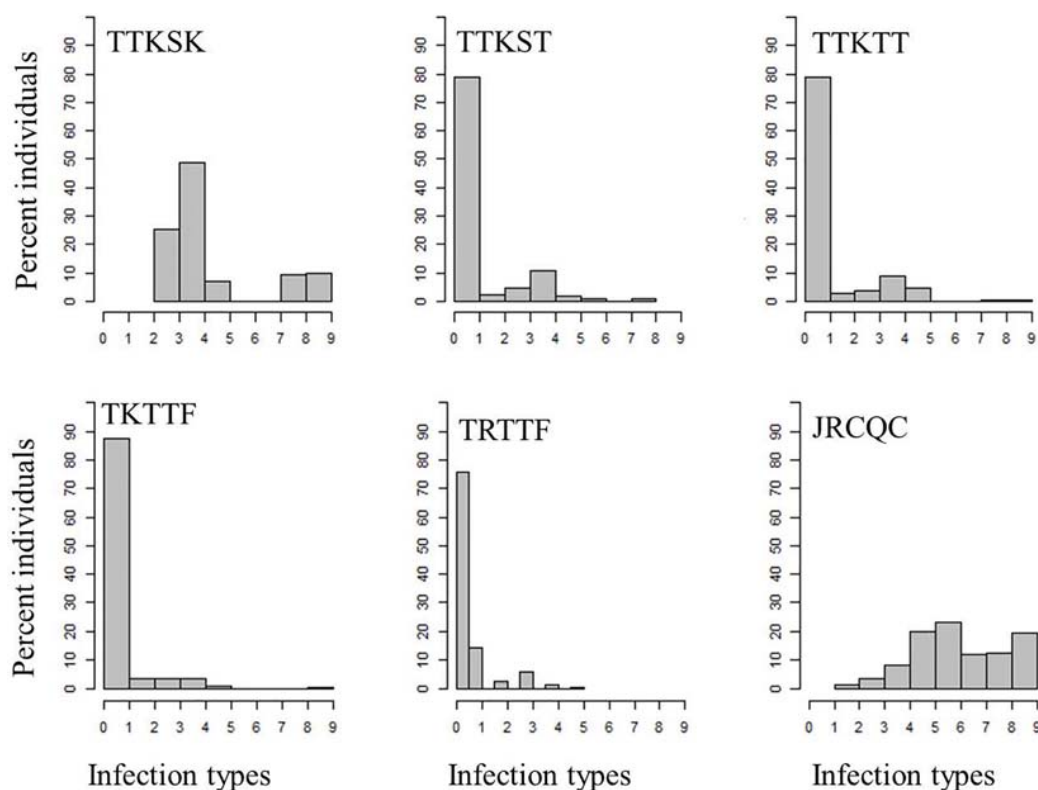


FIGURE 3 | Distributions of the seedling responses of the durum wheat genotypes to *Puccinia graminis* f. sp. *tritici* races TTKSK, TTKST, TTKTT, TKTTF, TRTTF, and JRCQC. X-axis corresponds to the linearized Stakman scale (0-to-9).

some of the durum genotypes (**Supplementary Figure 1**). This suggests intermediate familial relatedness between genotypes as described by Johnson et al. (2019). The PCA showed that the first two, three, four, and 10 PCs explained a cumulative variance of 9.4, 13.2, 16.5, and 31.4% of variation, respectively. The genotypes were clustered into three groups (Johnson et al., 2019) with majority of the lines grouped within the same cluster (**Supplementary Figure 2**). This is expected because the genotypes are from the same breeding program. Based on BIC values, mixed linear models that include both *Q* and *K* matrices were used for the GWAS for most traits. For traits associated with responses to *Pgt* races TTKST and TKTTF, the best GWAS regression models included the *K* matrix but not the *Q* matrix (**Table 1**).

Marker–Trait Associations

Association Analysis for Leaf Rust Response

The GWAS based on the linearized ITs to the three *Pt* isolates BBBSJ_TUN 20-1, EEEEE_ETH 13D14-1, and EEEEE_ETH 63-1 identified 64 significant SNPs (MTAs) at $FDR \leq 0.05$. Based on the LD between significant markers, these MTAs represented six loci located on chromosome arms 2AS, 2AL, 5BS, 6AL, and 6BL. The most significant marker/locus explained 6–31% of phenotypic variation (**Table 2**, **Figure 6**, and **Supplementary Table 6**). Chromosome arms 5BS and 6BL carried most of

the MTAs. Therefore, the pairwise LD between the significant markers on each of these chromosome arms were presented in **Supplementary Figure 3** that was used to determine the number of loci on chromosomes 5BS and 6BL.

On chromosome arm 2AS, the large-effect loci, *QLrdu.2AS* (Tag SNP: *IWB10489*, 67.5 cM, 61 Mbp) was associated with response to the Ethiopian isolates EEEEE_ETH 13D14-1 and EEEEE_ETH 63-1. On chromosome arm 2AL, *QLrdu.2AL* (*IWB38096*, 197.6 cM) was associated with response to race BBBSJ_TUN 20-1. On chromosome arm 5BS, two loci were associated with response to BBBSJ_TUN 20-1 and designated as *QLrdu.5BS-1* (*IWB47425*) and *QLrdu.5BS-2* (*IWB26157*). *QLrdu.5BS-1* explained higher phenotypic variation compared to *QLrdu.5BS-2*. These two loci spanned a genomic region from 2.0 to 35.8 cM corresponding to 4–21 Mbp on Svevo physical map (Maccaferri et al., 2019). On chromosome arm 6AL, a small-effect locus, *QLrdu.6AL* (*IWB24755*, 129.4 cM, 612 Mbp) was associated with response to EEEEE_ETH 63-1. An additional locus on chromosome arm 6BL, *QLrdu.6BL* (*IWB52926*, 154.6 cM, 696 Mbp) was also associated with response to EEEEE_ETH 63-1. All the leaf rust resistance loci identified in this study were race/isolate specific, except *QLrdu.2AS* that was associated with two Ethiopian isolates (**Table 2**, **Figure 6**, and **Supplementary Table 6**).

The postulation of the six *Lr* loci in each genotype in this germplasm was based on the most significant marker per locus

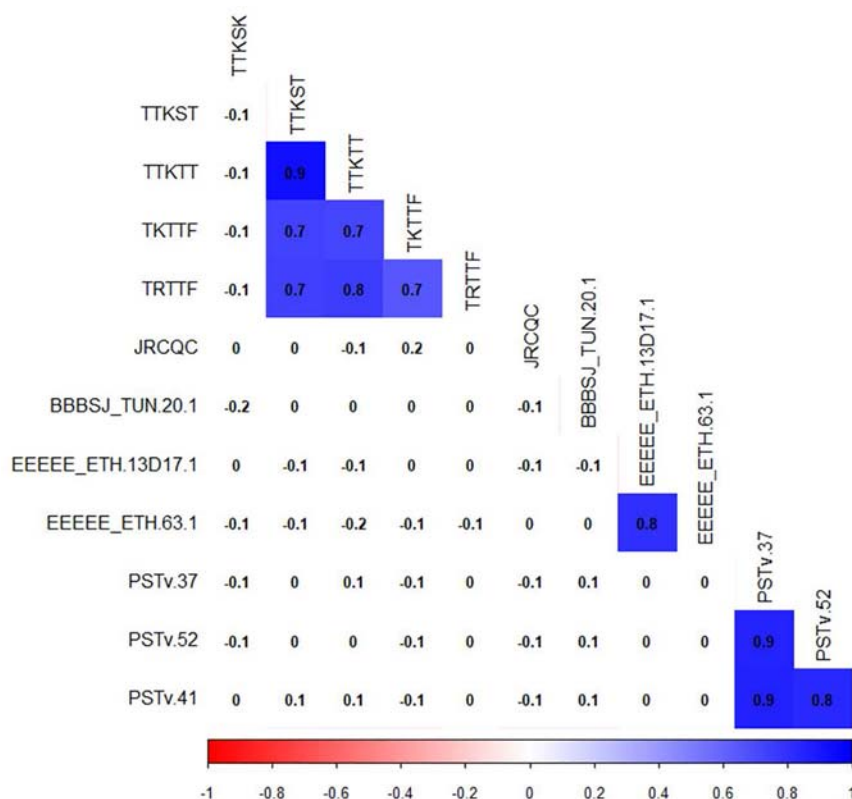


FIGURE 4 | Correlation between durum wheat line infection types to leaf rust, stem rust, and stripe rust pathogen races. Cells with significant correlations at P -value < 0.05 were in blue color.

and is presented in **Supplementary Table 1**. We found that all genotypes carry at least one of the identified loci in this study except lines D06707, D06710, D091721, and D97780. A total of 91% of the genotypes carry *QLrdu.6AL* and *QLrdu.6BL*, whereas 88% of the genotypes carry *QLrdu.2AS*, *QLrdu.6AL*, and *QLrdu.6BL*. Nine genotypes carry all the six identified loci in this study including Plaza (PI 613619), D98015, D98016, D01279, D011238, D03004, D05547, D101558, and D101650.

Association Analysis for Stripe Rust Response

The GWAS to the three *Pst* isolates PSTv-37, PSTv-52, and PSTv-41 identified 46 significant MTAs, corresponding to four loci located on chromosome arms 1BS, 5BL, and 7BL. The most significant SNP/locus explained 6–19% of phenotypic variation (**Table 3**, **Figure 7**, and **Supplementary Table 7**). Most of the MTAs were on chromosome arms 5BL and 7BL. Therefore, the pairwise LD between the significant markers on each of these chromosome arms were presented in **Supplementary Figure 4** that was used to determine the number of loci on each chromosome.

On chromosome arm 1BS, *QYrdu.1BS* (Tag SNP: *IWB31649*, 33 cM, 89 Mbp) was associated with response to race PSTv-52. On chromosome 5BL, two loci were detected. *QYrdu.5BL-1* (*IWA6271*, 187.1 cM, 682 Mbp) was associated with response to the three *Pst* races, whereas *QYrdu.5BL-2* (*IWB64287*, 193.4 cM,

691 Mbp) was associated with response to race PSTv-41. On chromosome 7BL, *QYrdu.7BL* (*IWB10533*, 187.5 cM, 697 Mbp) was associated with response to the three *Pst* races and explained most of the phenotypic variations. Two of the four identified stripe rust resistance loci in this study, *QYrdu.5BL-1* and *QYrdu.7BL* were associated with response to the three *Pst*-races, whereas the remaining *QYrdu.1BS* and *QYrdu.5BL-2* were race specific (**Table 3**, **Figure 7**, and **Supplementary Table 7**).

The postulation of the four *Yr* loci in each genotype in this germplasm was based on the most significant marker per locus and is presented in **Supplementary Table 2**. All genotypes carry at least one of the identified *Yr* loci in this study. A total of 78% of the genotypes carry *QYrdu.5BL-1* and *QYrdu.5BL-2*, whereas 52% of the genotypes carry *QYrdu.5BL-1* and *QYrdu.5BL-2* and *QYrdu.7BL*. Twenty-six genotypes carry all the four *Yr* loci identified in this study.

Association Analysis for Stem Rust Response

The GWAS detected 260 significant markers (MTAs), underlying 22 putative loci that were associated with stem rust response to the six *Pgt* races (TTKSK, TTKST, TTKTT, TKTTF, TRTTF, and JRCQC) (**Table 4**, **Figure 8**, and **Supplementary Table 8**). The highest number of MTAs were on chromosome arms 6AS (98 MTAs, three loci), 6AL (129 MTAs, three loci), 5AL (12 MTAs, three loci), and 6BL (seven MTAs, three loci). The pairwise LD

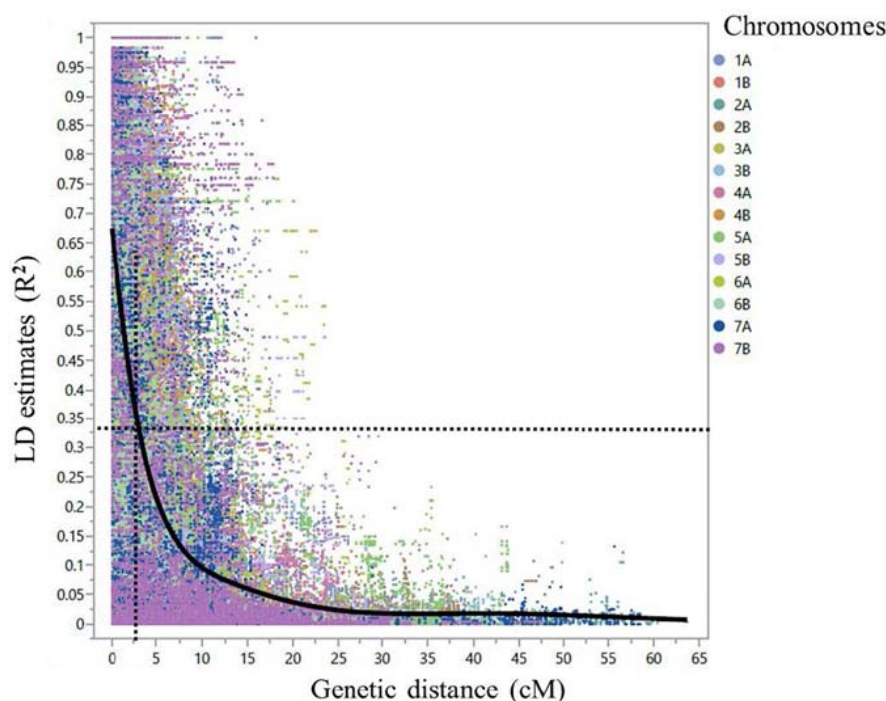


FIGURE 5 | Scatter plot demonstrating linkage disequilibrium (LD) decay across the 14 durum wheat chromosomes for the 248 durum wheat genotypes. The LD estimate (R^2) for pairs of SNPs was plotted against the corresponding genetic distance in centimorgan (cM) based on the tetraploid consensus map of Maccaferri et al. (2015). The dashed lines represent the LD decay that dropped by half at around 2.5 cM in average.

TABLE 1 | Bayesian Information Criterion of association mapping models for each trait.

| Rust trait | Races/isolates | Naive | Kinship | 2PCs + Kinship ^a | 3PCs + Kinship ^b | 4 PCs + Kinship ^c |
|-------------|------------------|--------|--------------|-----------------------------|-----------------------------|------------------------------|
| Leaf rust | BBBSJ_TUN 20-1 | 1206.9 | 1195.9 | 1183.5 | 1180.6 | 1175.6^d |
| | EEEE_ETH 13D17-1 | 683.5 | 672.4 | 671.1 | 661 | 658 |
| | EEEE_ETH 63-1 | 1088.5 | 1077.5 | 1074.1 | 1057.6 | 1049.3 |
| Stripe rust | PSTv-41 | 1204.8 | 1193.8 | 1180.9 | 1124 | 1173.3 |
| | PSTv-52 | 1158.3 | 1063.2 | 1131.1 | 1123.8 | 1120.5 |
| | PSTv-37 | 1180.6 | 1169.6 | 1159.2 | 1152.3 | 1150.8 |
| Stem rust | TTKSK | 1023.2 | 935.4 | 935.3 | 955.7 | 955.7 |
| | TTKST | 948.5 | 752.9 | 893.9 | 869.3 | 867.6 |
| | TKTt | 957.6 | 752.9 | 749.7 | 752.9 | 889.3 |
| | TKTTF | 796.5 | 666.5 | 702.5 | 686.4 | 685.0 |
| | TRTTF | 663.9 | 652.8 | 611.5 | 600.7 | 599.7 |
| | JRCQC | 1002 | 864.2 | 863.4 | 863.2 | 853.2 |

^a2PC, population structure matrix (Q matrix) based on the first two principal components explaining 9.4% of variation.

^b3PC, population structure matrix (Q matrix) based on the first three principal components explaining 13.2% of variation.

^c4PC, population structure matrix (Q matrix) based on the first four principal components explaining 16.5% of variation.

^dNumbers in bold indicate the lowest Bayesian Information Criterion that corresponds to the best regression model for each trait. The best model was used to investigate marker-trait associations.

between the significant markers on each of these chromosome arms were presented in **Supplementary Figure 5** and were used to determine the number of loci per chromosome. Other MTAs were identified on chromosomes 3AL (three MTAs, two loci), 4AL (four MTAs, a single locus), 5BL (two MTAs, two loci), and 7BL (two MTAs, a single locus). Each of the chromosome arms 1BL, 2BL, and 3BL carried a single MTA. Of the 22 identified loci, seven loci, *Qsrd.2BL*, *Qsrd.4AL*, *Qsrd.5AL-1*, *Qsrd.6AS-1*,

Qsrd.6AL-2, *Qsrd.6AL-3*, and *Qsrd.6BL-3*, were the most important loci in this study as they explained high phenotypic variations and/or associated with response to multiple *Pgt* races. These seven large-effect loci (highlighted in bold in **Table 4**) are the most robust *Sr* loci and were well represented in this germplasm ($MAF \geq 19\%$).

The most important large-effect locus identified on the distal end of chromosome arm 6AS was *Qsrd.6AS-1* (58–80

TABLE 2 | Summary of leaf rust resistance loci in the durum wheat genotypes.

| <i>P. triticina</i> race | Locus | Num. SNPs/locus ^a | Tag-SNP ^b | Chr. ^c | SNP alleles ^d | SNP major allele | SNP minor allele | MAF ^e | Position (cM) ^f | -Log10 (P-value) | R ^{2g} | pFDR ^h | Svevo genome v1 ⁱ | |
|--------------------------|-------------|------------------------------|----------------------|-------------------|--------------------------|------------------|------------------|------------------|----------------------------|------------------|-----------------|-------------------|------------------------------|-------------|
| | | | | | | | | | | | | | Start | End |
| BBBSL_TUN 20-1 | QLrdu.2AL | 1 | IWB38096 | 2A | T/C | T | C | 0.16 | 197.6 | 13.1 | 0.22 | 1.5E-10 | NA | NA |
| BBBSL_TUN 20-1 | QLrdu.5BS-1 | 26 | IWB47425 | 5B | A/C | A | C | 0.16 | 2.0 | 15.8 | 0.25 | 1.8E-12 | 5,077,659 | 5,077,759 |
| BBBSL_TUN 20-1 | QLrdu.5BS-2 | 19 | IWB26157 | 5B | A/G | G | A | 0.06 | 16.7 | 5.2 | 0.08 | 2.3E-03 | 21,105,871 | 21,105,971 |
| EEEE_ETH 13D14-1 | QLrdu.2AS | 2 | IWB10489 | 2A | T/C | C | T | 0.08 | 67.5 | 20.3 | 0.31 | 5.0E-17 | 61,159,123 | 61,159,023 |
| EEEE_ETH 63-1 | QLrdu.2AS | 2 | IWB10489 | 2A | T/C | C | T | 0.08 | 67.5 | 16.5 | 0.26 | 3.6E-13 | 61,159,123 | 61,159,023 |
| EEEE_ETH 63-1 | QLrdu.6AL | 1 | IWB24755 | 6A | T/C | C | T | 0.03 | 129.4 | 5.2 | 0.09 | 1.7E-02 | 612,235,063 | 612,235,163 |
| EEEE_ETH 63-1 | QLrdu.6BL | 13 | IWB52926 | 6B | A/G | G | T | 0.04 | 154.6 | 7.8 | 0.13 | 6.4E-05 | 695,708,680 | 695,708,580 |

^aDetails on all significant SNPs/locus were described in **Supplementary Table 6**.^bSignificant single-nucleotide polymorphism (SNP) with smallest marker-trait association P values per locus.^cChromosome arm of the locus.^dAlleles of the most significant marker in the loci. The underlined allele is the allele associated with resistance.^eMinor allele frequency of the most significant SNP/locus.^fSNP position based on the tetraploid consensus map of Maccaferri et al. (2015).^gThe proportion of phenotypic variation explained by the most significant SNP in the locus.^hP-value of the false discovery rate of the most significant SNP in the locus.ⁱPhysical position of SNP sequence based on the durum wheat genome sequence of Svevo available on International Wheat Genome Sequencing Consortium (Maccaferri et al., 2019).

MTAs, Tag-SNP: *IWB10558*, 0.2 cM, 2 Mbp) that was associated with response to races TTKST, TTKTT, TKTTF, and TRTTF but not to race TTKSK and JRCQC. *KASP_6AS_IWB10558* linked to the gene *Sr8155B1* was among the most significant markers in this locus. In addition, *KASP_6AS_IWB10558* was in high LD with other significant markers in *QSrdu.6AS-1* (Table 4, Supplementary Table 8, and Supplementary Figure 5) suggesting that the latter locus is indeed *Sr8155B1*. Two additional small-effect loci on chromosome 6AS and proximal to *QSrdu.6AS-1* were identified. *QSrdu.6AS-2* (*IWB67075*, 34.9 cM, 50 Mbp) was associated with response to race TTKST, whereas *QSrdu.6AS-3* (*IWA7295*, 45.9 cM, 86 Mbp) was associated with response to both races TTKST and TTKTT (Table 4, Figure 8, and Supplementary Table 8).

Two large-effect loci appeared on chromosome arm 6AL. *QSrdu.6AL-2* (Tag-SNP: *IWB69393*, 128.9 cM, 612 Mbp) was associated with response to race TTKSK, while *QSrdu.6AL-3* (Tag-SNP: *IWB41394*, 129.4 cM, 613 Mbp) was associated with response to race JRCQC. An additional small-effect locus on chromosome 6AL, *QSrdu.6AL-1* (Tag-SNP: *IWB31531*, 122.1 cM, 600 Mbp) was associated with response to race TKTTF. Even though, *QSrdu.6AL-2* and *QSrdu.6AL-3* were close based on their genetic positions (on tetraploid consensus map) and physical positions (Svevo genome v1), significant markers in these two loci were not in strong LD ($R^2 = 0.14$, Supplementary Figure 5). *QSrdu.6AL-2* and *QSrdu.6AL-3* appeared to be associated with *Sr13* gene/alleles. This is because *Sr13* diagnostic marker (*dCAPS_Sr13*) was among the most significant markers for race TTKSK and in LD with other significant SNPs in *QSrdu.6AL-2*. *Sr13* allele markers, *dCAPS_Sr13_R1cut* (identifying *Sr13a*-R1 allele) and *dCAPS_Sr13_R2nocut* (identifying *Sr13* R2 allele or *Sr13b*) were among significant markers for race JRCQC and in LD with significant SNPs in *QSrdu.6AL-3* (Table 4, Figure 8, and Supplementary Table 8).

The major allele ('T') of *IWB41394* that is the most significant SNP in *QSrdu.6AL-3* and present in 62 % of the durum genotypes was associated with susceptibility to race JRCQC. On the other hand, the most significant marker in *QSrdu.6AL-3* was *Sr13b* marker *dCAPS_Sr13_R2nocut*. The latter showed that 58% of the durum lines carry *Sr13b* associated with susceptibility to race JRCQC. Therefore, it is likely that the 'T' allele of *IWB41394* is associated with *Sr13b*. Overall, in 90.3% of the genotypes, there was agreement between marker *dCAPS_Sr13_R2nocut* and marker *IWB41394* in postulating *Sr13b* allele (Supplementary Table 9).

On chromosome arm 4AL, *QSrdu.4AL* (Tag-SNP: *IWA4651*, 162.4 cM, 719 Mbp) was another large-effect locus identified for response to race JRCQC. On chromosome arm 5AL, three loci were identified. *QSrdu.5AL-1* (*IWB62132*, 136.3 cM, 532 Mbp) was associated with response to multiple races TTKST, TTKTT, TRTTF, and TKTTF and explained 10–20% of phenotypic variation. In addition, two small-effect loci on chromosome 5AL, *QSrdu.5AL-2* (*IWB2075*, 183.0 cM, 623 Mbp) and *QSrdu.5AL-3* (*IWB14445*, 197.7 cM, 640 Mbp) were associated with response to race TKTTF. On chromosome arm 6BL, *QSrdu.6BL-1* (*IWB21973*, 103.7 cM, 622 Mbp) and *QSrdu.6BL-2* (*IWB5378*, 146.0 cM, 682 Mbp) was associated with response to race TRTTF

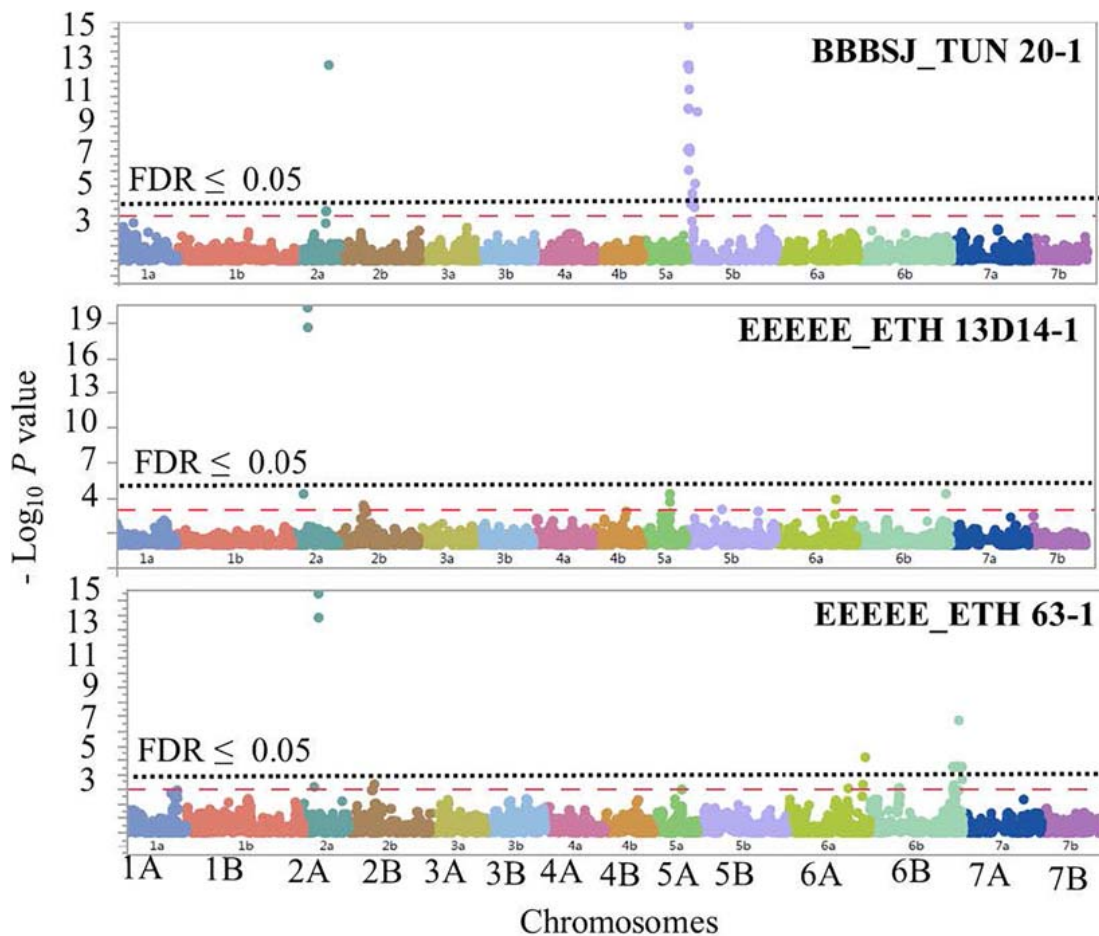


FIGURE 6 | Manhattan plots showing P -values for single-nucleotide polymorphism (SNP) markers associated with response to leaf rust in durum wheat genotypes to the three durum wheat type isolates BBBSJ_TUN 20-1, EEEEE_ETH 13D14-1, and EEEEE_ETH 63-1. The horizontal dashed red line indicates significance level at P -value ≤ 0.001 . The horizontal dotted black line indicates significance level at $FDR \leq 0.05$.

and TTKSK, respectively. *QSrdu.6BL-3* (*IWB46893*, 155.1 cM, 693 Mbp) was associated with responses to races TTKSK and JRCQC. The major allele of the most significant marker in *QSrdu.6BL-3*, *IWB46893*, was associated with resistance to TTKSK but with susceptibility to JRCQC.

On chromosome arm 5BL, two small-effect loci were identified: *QSrdu.5BL-1* (*IWB9652*, 181.5 cM, 675 Mbp) and *QSrdu.5BL-2* (*IWB64287*, 193.4 cM, 691 Mbp). Interestingly, *IWB64287* was also associated with response to *Pst*-race PSTv-41 (Tables 3, 4). This suggests that this locus on 5BL at 691 Mbp is associated with response to both stripe rust and stem rust and the allele 'C' of marker *IWB64287* provides resistance to both rust pathogens. Few MTAs were identified on each of the chromosomes 1BL, 2BL, 3AL, 3BL, and 7BL and most of these associations had minor effects on disease response (6–12%), except *QSrdu.2BL* (*IWB48212*, 193.6 cM, 789 Mbp) that explained relatively higher phenotypic variations (9–21%) to races TTKST, TTKTT, TKTTF, and TRTTF. Of the 22 identified loci for stem rust, five (*QSrdu.1BL*, *QSrdu.2BL*, *QSrdu.5AL-1*, *QSrdu.6AS-1*, and *QSrdu.6BL-3*) were associated with response

to more than one race while the remaining loci were race specific (Table 4, Figure 8, and Supplementary Table 8).

Frequencies of *Sr8155B1*, *Sr13*, and *Sr7b* in the Durum Wheat Genotypes and Their Marker Accuracies

Gene postulation for *Sr8155B1*, *Sr13* alleles, and *QSrdu.4AL* in each of the durum wheat genotypes is presented in Supplementary Tables 3, 5. Both phenotypic data (Supplementary Table 3) and marker data (Supplementary Table 9) were used to postulate the gene combinations present in each of the durum wheat genotypes. For the genotypic data, the markers *dCAPS_Sr13*, *dCAPS_Sr13_R1cut*, *dCAPS_Sr13_R2nocut*, *dCAPS_Sr13_R3nocut*, *KASP_6AS_IWB10558*, and *IWA4651* were used to postulate *Sr13a-R1*, *Sr13b*, *Sr13a-R3*, *Sr8155B1*, and *QSrdu.4AL* (designated in this study as *Sr7b*), respectively. We found that 81, 79, and 64% of the durum wheat genotypes carry *Sr13*, *Sr8155B1*, and *Sr7b*, respectively. A single breeding line (D07726) does not carry any of these genes.

TABLE 3 | Summary of stripe rust resistance loci in the durum wheat genotypes.

| Pst race | Locus | Num. SNPs/locus ^a | Tag-SNP ^b | Chr. ^c | SNP alleles ^d | SNP major allele | SNP minor allele | MAF ^e | Position (cM) ^f | -Log ₁₀ (P-value) | R ² _g | pFDR ^h | Svevo genome v1 ⁱ | |
|----------|-------------|------------------------------|----------------------|-------------------|--------------------------|------------------|------------------|------------------|----------------------------|------------------------------|-----------------------------|-------------------|------------------------------|-------------|
| | | | | | | | | | | | | | Start | End |
| PSTv-37 | QYrdv.5BL-1 | 7 | IWA6271 | 5B | A/G | G | A | 0.2 | 187.1 | 5.19 | 0.08 | 1.92E-03 | 681,513,107 | 681,512,935 |
| PSTv-37 | QYrdv.7BL | 33 | IWB10533 | 7B | T/C | C | T | 0.4 | 187.5 | 11.82 | 0.19 | 1.27E-09 | 696,924,344 | 696,924,444 |
| PSTv-52 | QYrdv.1BS | 1 | IWB31649 | 1B | A/G | A | G | 0.2 | 33.0 | 3.70 | 0.06 | 5.02E-02 | 88,957,108 | 88,957,007 |
| PSTv-52 | QYrdv.5BL-1 | 9 | IWA6271 | 5B | A/G | G | A | 0.2 | 187.1 | 4.79 | 0.08 | 4.94E-03 | 681,513,107 | 681,512,935 |
| PSTv-52 | QYrdv.7BL | 33 | IWB10533 | 7B | T/C | C | T | 0.4 | 187.5 | 10.33 | 0.17 | 2.91E-08 | 696,924,344 | 696,924,444 |
| PSTv-41 | QYrdv.5BL-1 | 9 | IWA6271 | 5B | A/G | G | A | 0.2 | 187.1 | 6.21 | 0.10 | 1.92E-04 | 681,513,107 | 681,512,935 |
| PSTv-41 | QYrdv.5BL-2 | 1 | IWB64287 | 5B | A/C | C | A | 0.1 | 193.4 | 4.49 | 0.07 | 9.30E-03 | 691,154,062 | 691,153,962 |
| PSTv-41 | QYrdv.7BL | 33 | IWB10533 | 7B | T/C | C | T | 0.4 | 187.5 | 10.26 | 0.17 | 3.35E-08 | 696,924,344 | 696,924,444 |

^aDetails on all significant SNPs/locus were described in **Supplementary Table 7**.^bSignificant single-nucleotide polymorphism (SNP) with smallest marker-trait association P-values per locus.^cChromosome arm of the locus.^dAlleles of the most significant marker in the locus. The underlined allele is the allele associated with resistance.^eMinor allele frequency of the most significant SNP/locus.^fSNP position based on the tetraploid consensus map of Maccaferri et al. (2015).^gThe proportion of phenotypic variation explained by the most significant SNP in the locus.^hP-value of the false discovery rate of the most significant SNP in the locus.ⁱPhysical position of SNP sequence based on the durum wheat genome sequence of Svevo available on International Wheat Genome Sequencing Consortium (Maccaferri et al., 2019).

A total of 61% of the durum genotypes carry an *Sr13* allele and *Sr8155B1*, whereas 50% of the durum collection carry an *Sr13* allele and *Sr7b*. We found that 54% of the durum genotypes have a least *Sr8155B1* and *Sr7b* and 40% of the genotypes have the three genes *Sr13*, *Sr8155B1*, and *Sr7b*. Based on *Sr13* allele markers, *Sr13* functional alleles *Sr13a-R1*, *Sr13b*, and *Sr13a-R3* were identified in the durum genotypes. *Sr13b* was the most common allele, being present in 56% of the durum genotypes. *Sr13a-R1* and *Sr13a-R3* were less frequent and occurred in only 17 and 7% of the durum accessions, respectively (**Supplementary Table 9**).

Because gene postulation for these three genes was possible based only on the phenotype, we determined the accuracies of markers *dCAPS_Sr13*, *IWB69393*, *KASP_6AS_IWB10558*, and *IWA4651*. For the gene *Sr13*, the accuracy for *dCAPS_Sr13* and *IWB69393* was 100 and 95% (3% false positives and 2% false negatives), respectively. For *Sr8155B1*, the marker *KASP_6AS_IWB10558* had an accuracy of 99.6% (0.4% false positives), whereas for *Sr7b*, the marker *IWA4651* had an accuracy of 98.8% (1.2% false positives). The postulation of the remaining three large-effect *Sr* loci in each genotype (**Supplementary Table 3**) showed that 30 genotypes carry *Sr8155B1*, *Sr13*, *Sr7b*, *QSrdu.2BL*, *QSrdu.5AL-1*, and *QSrdu.6BL-3*.

DISCUSSION

Leaf Rust Resistance in Durum Wheat Genotypes

All the durum genotypes were resistant to the common wheat type race MBDS5 that is widely distributed in the wheat growing regions of the United States (Kolmer and Hughes, 2014). This agrees with previous studies indicating that *Pt*-isolates from common wheat are generally avirulent on durum wheat (Singh, 1991; Huerta-Espino and Roelfs, 1992; Ordoñez and Kolmer, 2007a; Aoun et al., 2016). Herrera-Foessel et al. (2014) reported that most of the CIMMYT durum wheat germplasm carry *Lr72* that is effective against common wheat type races. Thus, *Lr72* could be also present in the durum wheat genotypes in this study. Many of the genotypes in our study were susceptible to Mexican, Moroccan, Tunisian, and Ethiopian durum wheat type isolates. None of the durum genotypes were resistant to the *Pt*-Mexican race BBBQJ. The latter is similar to a race collected on durum wheat in California (Kolmer, 2013) and on hard red winter wheat in Kansas (Kolmer, 2015b). Even though *Pt*-race BBBQJ is not yet present in North Dakota, introgression of leaf rust resistance to this race in the NDSU durum wheat lines will help the growers in tackling in future challenges. For instance, previously identified *Lr* genes like those identified in CIMMYT germplasm (Herrera-Foessel et al., 2007, 2008a,b; Huerta-Espino et al., 2009) and in the USDA–National Small Grains Collection (NSGC) of durum wheat (Aoun et al., 2016, 2017, 2019) could be used to enhance leaf rust resistance to race BBBQJ in the NDSU durum wheat germplasm. The Ethiopian isolates of race EEEEE were virulent to only 10–28% of the durum genotypes. Even though, the two Ethiopian isolates in this study carry the same race (EEEE) on Thatcher wheat differentials, there were differences in their

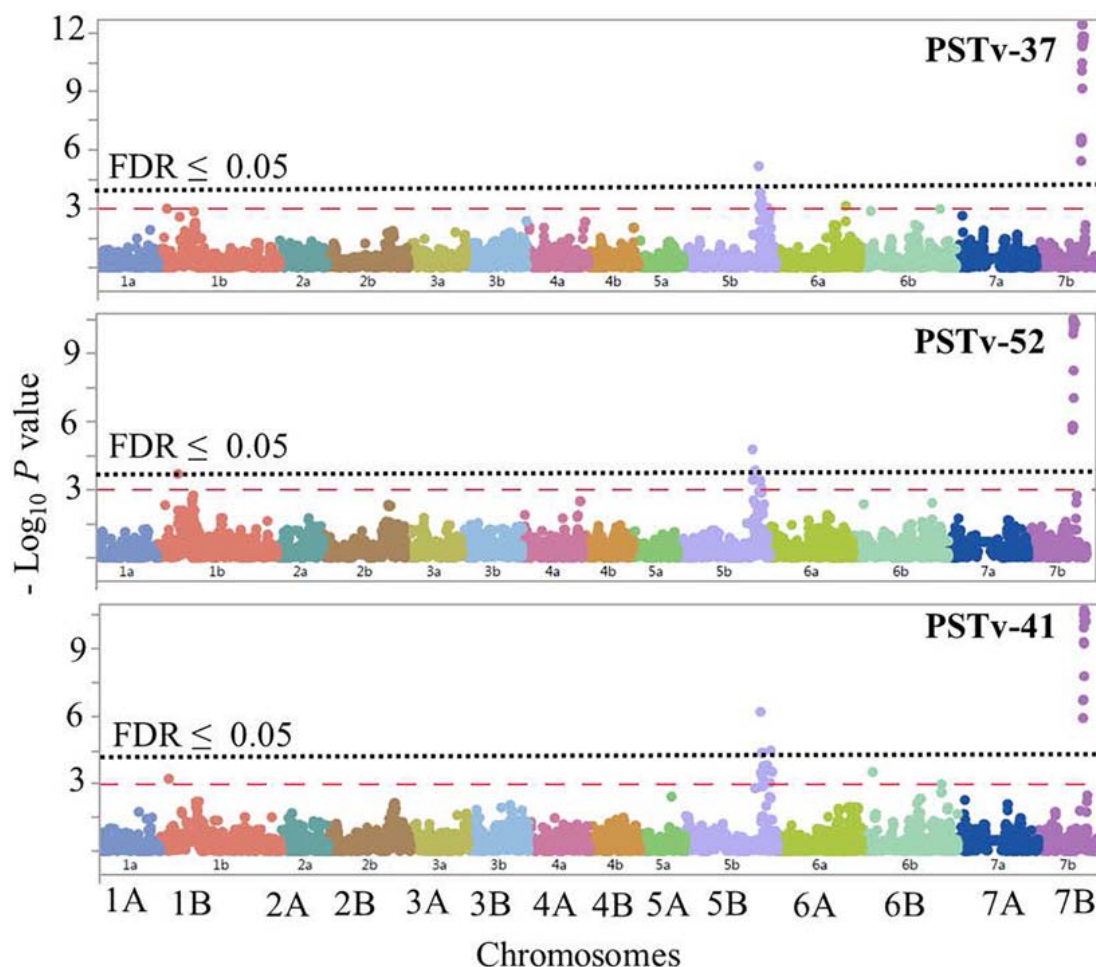


FIGURE 7 | Manhattan plots showing P -values for single-nucleotide polymorphism (SNP) markers associated with response to stripe rust in durum wheat genotypes to the three *Pst* isolates PSTv-37, PSTv-52, and PSTv-41. The horizontal dashed red line indicates significance level at P -value ≤ 0.001 . The horizontal dotted black line indicates significance level at $FDR \leq 0.05$.

virulence profiles on durum wheat genotypes in our study. These results agree with Aoun et al. (2020) observations showing that different virulence phenotypes were found within a collection of isolates of race EEEEE based on a set of durum wheat differentials.

Comparative mapping between the identified six all-stage leaf rust resistance loci in this study and designated wheat *Lr* genes showed that any of the two loci on chromosome 5BS could be *Lr52* that was previously identified in the durum wheat cultivar Wallaroi (Singh et al., 2010). Similarly, *QLrdu.6AL* is most likely *Lr64* that originated from wild emmer wheat (*Triticum dicoccoides*) (Dyck, 1994; McIntosh et al., 2009; Kolmer et al., 2019). The remaining loci did not map close to known *Lr* genes and thus could be novel. Comparison of the map locations suggests that *QLrdu.2AS* (67.5 cM, 61 Mbp) is likely the same locus which was earlier found associated with leaf rust response in durum wheat and tagged by the SSR marker *wmc522* (63.6 cM, 58 Mbp) (Maccafferri et al., 2010). The nine genotypes that carry all the six identified *Lr* loci in this study are useful to keep these resistance sources in future released varieties.

Stripe Rust Resistance in Durum Wheat Genotypes

Many of the durum wheat genotypes (67–69%) in this study were resistant to the three U.S. *Pst* races (PSTv-37, PSTv-52, and PSTv-41). A previous study that screened a worldwide collection of elite durum wheat lines to six US and Italian *Pst*-races (including PSTv-37) showed that only 7.8–31.5% of the genotypes were resistant (Liu et al., 2017). This suggests that the durum wheat collection in this study had undergone selection to accumulate potentially useful loci for stripe rust resistance to the North American *Pst* races. The durum wheat responses to these three *Pst*-races used in this study were highly correlated, showing that the NDSU durum genotypes had a broad spectrum of stripe rust resistance.

With rapid and dangerous shifts in *Pst* populations globally (Solh et al., 2012), our study will help durum wheat breeding programs by providing new stripe rust resistance sources. We identified four loci associated with all-stage stripe rust resistance that did not correspond to any designated stripe rust resistance

TABLE 4 | Summary of stem rust resistance loci in the durum wheat genotypes.

| Pgt race | Locus ^a | Num. SNPs/locus ^b | Tag-SNP ^c | Chr. ^d | SNP alleles ^e | SNP major allele | SNP minor allele | MAF ^f | Position (cM) ^g | -Log ₁₀ (P-value) | R ^{2h} | pFDR ⁱ | Svevo genome v1 ^j | |
|----------|--------------------|---------------------------------|--------------------------------|-------------------|-----------------------------|---------------------|---------------------|------------------|-------------------------------|---------------------------------|-----------------|-------------------|------------------------------|-------------|
| | | | | | | | | | | | | | Start | End |
| TTKSK | QSrd.3AL-2 | 1 | IWB72044 | 3A | A/G | A | G | 0.05 | 177.9 | 3.85 | 0.06 | 1.64E-02 | 736,648,215 | 736,64,8115 |
| TTKSK | QSrd.3BL | 1 | IWB49397 | 3B | T/G | T | G | 0.05 | 77.1 | 7.61 | 0.12 | 4.73E-06 | 370,964,536 | 370,964,636 |
| TTKSK | QSrd.5BL-1 | 1 | IWB9652 | 5B | T/C | C | T | 0.08 | 181.5 | 3.28 | 0.05 | 5.42E-02 | 674,697,988 | 674,698,088 |
| TTKSK | QSrd.5BL-2 | 1 | IWB64287 | 5B | A/C | C | A | 0.07 | 193.4 | 3.36 | 0.05 | 4.58E-02 | 691,154,062 | 691,153,962 |
| TTKSK | QSrd.6AL-2 | 92 | IWB69393/ dCAPS_Sr13 | 6A | T/C | T | C | 0.19 | 128.9 | 19.92 | 0.32 | 6.57E-17 | 611,710,729 | 611,710,829 |
| TTKSK | QSrd.6BL-2 | 6 | IWB5378 | 6B | T/G | G | T | 0.05 | 146.0 | 7.61 | 0.12 | 4.73E-06 | 682,240,129 | 682,240,229 |
| TTKSK | QSrd.6BL-3 | 1 | IWB46893 | 6B | A/G | G | A | 0.38 | 155.1 | 3.46 | 0.05 | 3.65E-02 | 693,337,728 | 693,337,628 |
| TTKST | QSrd.1BL | 1 | IWB50554 | 1B | A/G | G | A | 0.11 | 27.6 | 5.66 | 0.09 | 3.28E-04 | NA | NA |
| TTKST | QSrd.2BL | 1 | IWB48212 | 2B | A/C | A | C | 0.20 | 193.6 | 12.77 | 0.20 | 4.11E-11 | 789,417,490 | 789,417,417 |
| TTKST | QSrd.5AL-1 | 1 | IWB62132 | 5A | T/G | G | T | 0.21 | 136.3 | 11.84 | 0.18 | 3.25E-10 | 532,077,979 | 532,077,878 |
| TTKST | QSrd.6AS-1 | 80 | IWB10558/ KASP_6AS_IWB10558 | 6A | T/C | C | T | 0.20 | 0.2 | 14.71 | 0.23 | 1.26E-11 | 1,590,026 | 1,590,126 |
| TTKST | QSrd.6AS-2 | 3 | IWB67075 | 6A | A/G | A | G | 0.09 | 34.9 | 3.41 | 0.05 | 4.99E-02 | 50,134,208 | 50,134,274 |
| TTKST | QSrd.6AS-3 | 1 | IWA7295 | 6A | T/G | T | G | 0.03 | 45.9 | 3.67 | 0.05 | 2.76E-02 | 86,025,214 | 86,025,359 |
| TTKTT | QSrd.1BL | 1 | IWB50554 | 1B | A/G | G | A | 0.11 | 27.6 | 5.28 | 0.08 | 7.78E-04 | NA | NA |
| TTKTT | QSrd.2BL | 1 | IWB48212 | 2B | A/C | A | C | 0.20 | 193.6 | 13.49 | 0.21 | 7.67E-12 | 789,417,490 | 789,417,417 |
| TTKTT | QSrd.3AL-1 | 2 | IWB36155 | 3A | T/C | T | C | 0.04 | 90.4 | 3.91 | 0.06 | 1.45E-02 | 572,456,904 | 572,456,785 |
| TTKTT | QSrd.5AL-1 | 1 | IWB62132 | 5A | T/G | G | T | 0.21 | 136.3 | 12.61 | 0.20 | 5.60E-11 | 532,077,979 | 532,077,878 |
| TTKTT | QSrd.6AS-1 | 80 | IWA5416/ KASP_6AS_ IWB10558 | 6A | T/C | T | C | 0.21 | 0.2 | 16.62 | 0.26 | 2.62E-13 | 1,198,024 | 1,197,947 |
| TTKTT | QSrd.6AS-3 | 10 | IWA7295 | 6A | T/G | T | G | 0.03 | 45.9 | 4.21 | 0.06 | 8.39E-03 | 86,025,214 | 86,025,359 |
| TKTTF | QSrd.2BL | 1 | IWB48212 | 2B | A/C | A | C | 0.20 | 193.6 | 8.36 | 0.13 | 1.40E-06 | 789,417,490 | 789,417,417 |
| TKTTF | QSrd.5AL-1 | 1 | IWB62132 | 5A | T/G | G | T | 0.21 | 136.3 | 6.42 | 0.10 | 6.58E-05 | 532,077,979 | 532,077,878 |
| TKTTF | QSrd.5AL-2 | 10 | IWB2075 | 5A | A/G | A | G | 0.03 | 183.0 | 4.72 | 0.07 | 2.61E-03 | 623,114,829 | 623,114,760 |
| TKTTF | QSrd.5AL-3 | 1 | IWB14445 | 5A | T/G | G | T | 0.04 | 197.7 | 3.60 | 0.05 | 3.04E-02 | 640,125,144 | 640,125,045 |
| TKTTF | QSrd.6AS-1 | 74 | IWB60233/ KASP_6AS_IWB10558 | 6A | T/C | T | C | 0.12 | 0.9 | 10.83 | 0.19 | 1.60E-07 | 3,721,352 | 3,721,450 |
| TKTTF | QSrd.6AL-1 | 3 | IWB31531 | 6A | A/G | A | G | 0.08 | 122.1 | 3.75 | 0.06 | 2.21E-02 | 600,285,732 | 600,285,802 |
| TRTTF | QSrd.2BL | 1 | IWB48212 | 2B | A/C | A | C | 0.20 | 193.6 | 5.69 | 0.09 | 5.84E-04 | 789,417,490 | 789,417,417 |
| TRTTF | QSrd.5AL-1 | 1 | IWB62132 | 5A | T/G | G | T | 0.21 | 136.3 | 7.40 | 0.12 | 1.09E-04 | 532,077,979 | 532,077,878 |
| TRTTF | QSrd.6AS-1 | 58 | IWB53754/ KASP_6AS_IWB10558 | 6A | A/G | G | A | 0.21 | 0.2 | 8.34 | 0.13 | 3.58E-05 | 1,202,823 | 1,202,923 |
| TRTTF | QSrd.6BL-1 | 5 | IWB21973 | 6B | A/G | A | G | 0.16 | 103.7 | 4.84 | 0.08 | 2.78E-03 | 621,527,086 | 621,527,186 |
| TRTTF | QSrd.7BL | 2 | IWB17567 | 7B | T/G | G | T | 0.05 | 147.0 | 4.40 | 0.07 | 7.19E-03 | 675,357,404 | 675,357,554 |
| JRCQC | QSrd.4AL | 4 | IWA4651 | 4A | A/G | A | G | 0.33 | 162.4 | 7.20 | 0.11 | 1.70E-04 | 718,619,698 | 718,619,565 |

(Continued)

TABLE 4 | Continued

| Pgt race | Locus ^a | Num. SNPs/locus ^b | Tag-SNP ^c | Chr. ^d | SNP alleles ^e | SNP major allele | SNP minor allele | MAF ^f | Position (cM) ^g | -Log ₁₀ (P-value) | R ^{2h} | pFDR ⁱ | Svevo genome v1 ^j | |
|----------|--------------------|------------------------------|---|-------------------|--------------------------|------------------|------------------|------------------|----------------------------|------------------------------|-----------------|-------------------|------------------------------|-------------|
| | | | | | | | | | | | | | Start | End |
| JRCQC | QSrdu.6AL-3 | 39 | IWB41394/ dCAPS_Sr13_2 R2nocut/ dCAPS_Sr13_R1cut | 6A | T/C | T | C | 0.38 | 129.4 | 5.83 | 0.09 | 1.16E-03 | 613,212,821 | 613,212,744 |
| JRCQC | QSrdu.6BL-3 | 1 | IWB46893 | 6B | A/G | G | A | 0.38 | 155.1 | 5.83 | 0.09 | 1.16E-03 | 693,337,728 | 693,337,628 |

^aThe loci in bold represent the most important stem rust resistance loci identified in this study.

^bDetails on all significant SNPs/locus were described in **Supplementary Table 7**.

^cSignificant single-nucleotide polymorphism (SNP) with smallest marker-trait association P values per locus.

^dChromosome arm of the locus.

^eAlleles of the most significant marker in the locus. The underlined allele is the allele associated with resistance.

^fMinor allele frequency of the most significant SNP/locus.

^gSNP position based on the tetraploid consensus map of Maccaferri et al. (2015).

^hThe proportion of phenotypic variation explained by the most significant SNP in the locus.

ⁱP-value of the false discovery rate of the most significant SNP in the locus.

^jPhysical position of SNP sequence based on the durum wheat genome sequence of Svevo available on International Wheat Genome Sequencing Consortium (Maccaferri et al., 2019).

genes. At the same time, some of the loci identified in this study were mapped close to not yet characterized stripe rust resistance quantitative trait loci (QTL) in the literature. For instance, *QYrdu.1BS* (*IWB31649*, 33.0 cM) was located close to previously identified locus *Yrdurum-1BS.1* (34.1–40.1 cM) that was associated with stripe rust response in a worldwide collection of elite durum wheat (Liu et al., 2017). Similarly, the *QYrdu.5BL-1* (*IWA6271*, 187.1 cM, 682 Mbp) was mapped close to the stripe rust resistance QTL, *QYr.usw-5B* (*IWA7066*, 179.6 cM, 674 Mbp, Lin et al., 2018) that was earlier detected in the durum wheat line W9262-260D3 (Kyle*2/Biodur). The position of *QYrdu.7BL* (*IWB10533*, 187.5 cM, 697 Mbp) also overlaps with that of *Yrdurum-7BL* (184.5–190.5 cM) that was associated with stripe rust response at seedling stage in elite durum wheat genotypes (Liu et al., 2017). At the similar location, Lin et al. (2018) also identified *QYr.usw-7B* (181.1 cM, 694 Mbp) in the durum wheat line W9262-260D3 (Kyle*2/Biodur) to Canadian isolates at seedling stage and to Mexican races at adult-plant stage. Further research warrants to characterize the four stripe rust resistance loci detected in this study and study their relationship with those previously identified in the literature. The 26 genotypes that carry all the four *Yr* loci identified in this study are excellent sources to introgress these stripe rust resistance sources in future durum wheat varieties.

Stem Rust Resistance in Durum Wheat Genotypes

The majority of durum wheat genotypes were resistant to the three Ug99-lineage races TTKSK, TTKST, and TTKTT. Interestingly, 19% of the genotypes were susceptible to race TTKSK while only 1% of the genotypes were susceptible to the other two Ug-99 lineage races TTKST and TTKTT. This suggests that these durum advanced breeding lines carry stem rust resistance gene(s)/allele(s), such as *Sr8155B1*, that are effective against TTKST and TTKTT but ineffective against TTKSK. Therefore, a combination of multiple *Sr* genes in the newly developed durum wheat cultivars is recommended for effective resistance to different races of the Ug99 lineage. Similarly, only one line was susceptible to the Digalu race (TKTTF) (Olivera et al., 2015). The durum genotypes were all resistant to race TRTTF. In contrast to TRTTF, race JRCQC that is adapted to durum wheat (Hundie et al., 2019) in Ethiopia was the most virulent race on the durum genotypes in our study. This suggests that *Sr* genes/alleles effective to races TTKSK, TTKST, TTKTT, TKTTF, and TRTTF do not provide resistance to JRCQC. Olivera et al. (2015) showed that races JRCQC, TRTTF, and TKTTF are phylogenetically different from Ug99-lineage races. Therefore, *Sr* genes effective to each of these race lineages could be different. This implies that a combination of diverse *Sr* genes should be implemented in newly released cultivars.

The durum wheat genotypes in this study showed higher levels of stem rust resistance compared to germplasm collections used in previous studies. For example, in a durum wheat collection from different durum wheat-growing regions in Mediterranean countries, the Southwestern United States, and Mexico, 42.1, 18.6, and 52.5% of the tested accessions were susceptible to

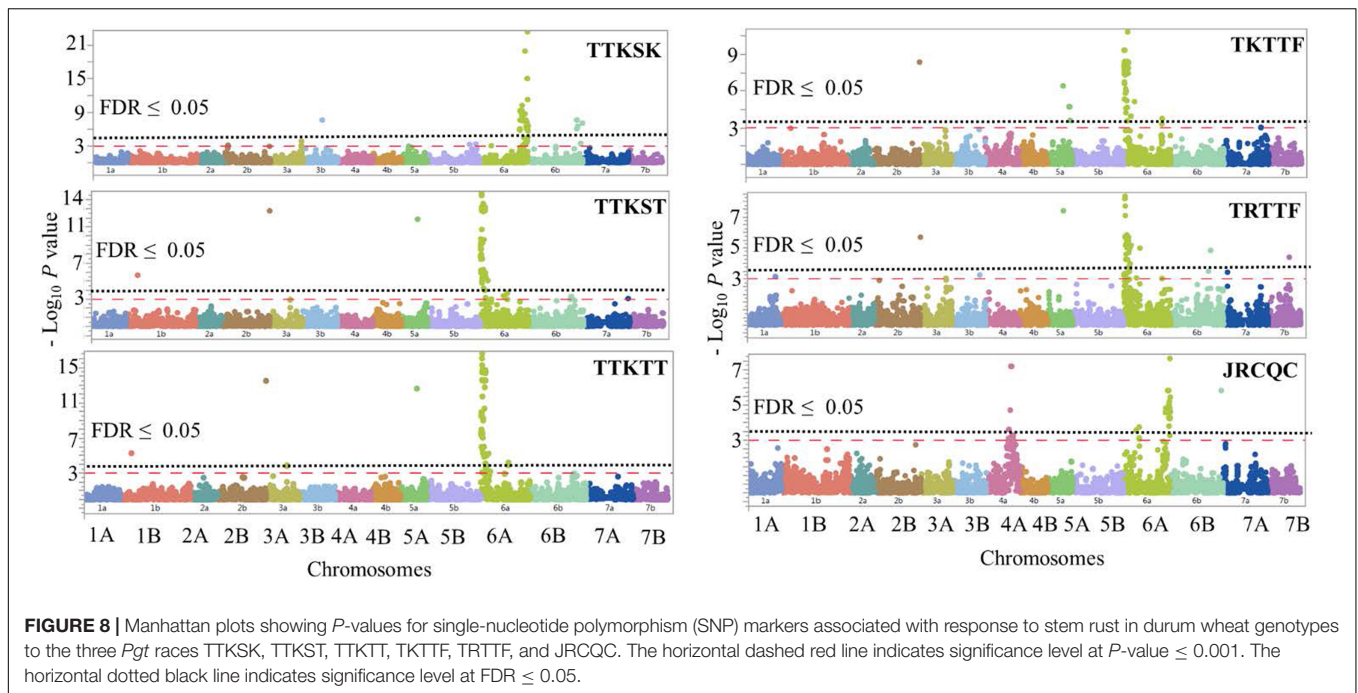


FIGURE 8 | Manhattan plots showing P -values for single-nucleotide polymorphism (SNP) markers associated with response to stem rust in durum wheat genotypes to the three *Pgt* races TTKSK, TTKST, TTKTT, TKTTF, TRTTF, and JRCQC. The horizontal dashed red line indicates significance level at P -value ≤ 0.001 . The horizontal dotted black line indicates significance level at $FDR \leq 0.05$.

TTKSK, TRTTF, and JRCQC, respectively (Letta et al., 2014). In another study (Chao et al., 2017), most of the USDA–NSGC of durum wheat collection comprised of landraces, breeding lines, and cultivars were found susceptible to TTKSK (81.6%), TRTTF (72.1%), and JRCQC (90.6%). This shows that the NDSU breeding program selected for stem rust resistance to most of the *Pgt*-races used in this study. It was reported that resistance to the Ug99 lineage in the North American durum cultivars is mainly due to *Sr13* alleles that were first identified in durum wheat and was then transferred to hexaploid wheat (Knott, 1990). However, in our study we observed variations in the ITs to the *Pgt*-races. For instance, the most common resistant IT to races TTKST, TTKTT, TKTTF, TRTTF (IT = 0), indicative of *Sr8155B1*, was much lower compared to the most common resistant infection type to TTKSK (IT = 2–) and JRCQC (IT = 22+). This suggests that stem rust genetic architecture in this durum wheat collection is much more complex and multiple genes/alleles could be identified in this durum germplasm. In this germplasm, we found that 40% of the durum genotypes carry *Sr13*, *Sr8155B1*, and *Sr7b* and 30 genotypes (12%) carry large-effect loci identified in this study including *Sr8155B1*, *Sr13a/Sr13b*, *Sr7b*, *QSRdu.2BL*, *QSRdu.5AL-1*, and *QSRdu.6BL-3*. This gene/loci combination is critical to keep in future released durum wheat varieties. The remaining 15 *Sr* loci that explained low phenotypic variation or associated with relatively low MAF need to be first validated before being used in breeding programs.

Comparative mapping showed that out of the 22 identified all-stage stem rust resistance loci in this study, four loci corresponded to cataloged *Sr* genes/alleles. In addition, eight loci in this study were mapped close to previously detected stem rust resistance QTL that were not yet cataloged in wheat. *QSRdu.1BL* was also found close to the DArT marker *wPt-1876* (26.3 cM)

that was associated with stem rust response in durum wheat (Letta et al., 2014). The locus *QSRdu.2BL* (*IWB48212*, 193.6 cM, 789 Mbp) was mapped close to SSR marker *wmc356* (788 Mbp) that has been found associated with stem rust response in durum wheat (Letta et al., 2014). Within the genomic regions of *QSRdu.3AL-1* (*IWB36155*, 90.4 cM, 572 Mbp) and *QSRdu.3AL-2* (*IWB72044*, 177.9 cM, 737 Mbp), Letta et al. (2013) identified two stem rust resistance loci in durum wheat tagged with the SSR marker *wmc428* (93.8 cM, 589 Mbp) and DArT marker (*wPt-8203*, 178.3 cM). The locus *QSRdu.4AL* (*IWA4651*, 162.4 cM, 719 Mbp) that was associated with response to race JRCQC was close to the mapping position of *Sr7* locus (McIntosh et al., 1995; Saini et al., 2018) and it is likely *Sr7b*. We found that 64% of the durum genotypes carry *Sr7b* and it is important to keep it in future released varieties, especially that only few known genes confer resistance to race JRCQC. Within the genomic region of *QSRdu.4AL* (*Sr7b*), Letta et al. (2014) identified a locus tagged with the SSR marker *barc78* (161.7 cM, 656 Mbp) associated with response to race JRCQC at seedling stage in elite durum wheat panel. In the same durum wheat panel, Letta et al. (2013) identified two MTAs on chromosome arm 4AL tagged by the DArT markers *wPt-9196* (157.7 cM) and *wPt-0798* (161.7 cM) associated with stem rust response at adult-plant stage in field trials in Ethiopia. Proximal to the genomic region of *QSRdu.5AL-1*, a MTA represented with the SSR marker *gwm1570* (134.5 cM) was associated with stem rust seedling response in durum wheat (Letta et al., 2014). Similarly, the genomic region near *QSRdu.5AL-2* and *QSRdu.5AL-3* were found to carry two stem rust resistance loci tagged with markers *gwm126* (191.2 cM) and *gwm291* (205.0 cM) in durum wheat in field trials in Ethiopia (Letta et al., 2013). On chromosome 5BL and at a close genomic region to *QSRdu.5BL-2*, Letta et al. (2014) detected a GWAS hit

tagged by DArt marker *wPt-0566* (191.6 cM) associated with stem rust seedling response in durum wheat (Letta et al., 2014).

The locus *QSrdu.6AS-1* (KASP_6AS_IWB10558, 0.2 cM, 2 Mbp) that was associated with resistance to race TTKST, TTKTT, TKTTF, and TRTTF was identified in the region of *Sr8155B1*. This gene was first identified in the durum wheat line 8155-B1 and known to confer resistance against race TTKST (Nirmala et al., 2017). The gene *Sr8155B1* was later reported in the durum wheat cultivar 'Lebsock' and provided resistance to race TRTTF (Saini et al., 2018). In our study, we observed that *Sr8155B1* provides resistance to additional *Pgt*-races TTKTT and TKTTF. In agreement with Nirmala et al. (2017), we found that this gene is common in the Midwestern durum wheat with 79% of the breeding lines and cultivars carrying this gene. Based on *Sr13* diagnostic markers, *QSrdu.6AL-2* and *QSrdu.6AL-3* were found to be associated with *Sr13* gene/alleles. *Sr13* is known to be common in North American and CIMMYT durum wheat cultivars (Jin, 2005; Singh et al., 2015) and is present in 84% of this durum wheat germplasm. *Sr13a* that confers resistance to JRCQC is present in only 17% of the durum genotypes in this study. However, 66% of the genotypes were resistant to JRCQC. This is most likely explained by the presence of other genes conferring resistance to JRCQC, e.g., *Sr7b*. *Sr13* gene/allele CAPS markers used in this study are difficult to be used in high-throughput genotyping for marker assisted selection. Therefore, the most significant SNPs in *QSrdu.6AL-2* (e.g., IWB69393) and *QSrdu.6AL-3* (e.g., IWB41394) can be converted into KASP or thermal asymmetric reverse PCR (STARP) markers to postulate the presence of *Sr13* gene and *Sr13b* allele, respectively.

The locus *QSrdu.7BL* (IWB17567, 147.0 cM, 675 Mbp) that was associated with response to race TRTTF is mapped close to the gene *Sr17*. The gene *Sr17* has been reported in tetraploid wheat and synthetic bread wheat (Bansal et al., 2008). However, race TRTTF is virulent to *Sr17*, therefore *QSrdu.7BL* is likely linked to *Sr17* or a new allele of *Sr17*. Close to the genomic region of *QSrdu.7BL*, Letta et al. (2013) also reported a stem rust resistance locus in durum wheat tagged by DArt marker *wPt-8615* (154.0 cM).

CONCLUSION

We investigated the levels of all-stage resistance in durum wheat genotypes adapted to the Midwest region of the U.S. against six *Pt*-races, three *Pst*-races, and six *Pgt*-races. Many of the durum wheat breeding lines and cultivars were susceptible to durum wheat type *Pt* isolates, whereas all lines were resistant to the common wheat type *Pt* isolate. In contrast to leaf rust, many of the durum wheat genotypes has high levels of resistance to most stripe rust and stem rust pathogen races. Association mapping revealed six leaf rust resistance loci located on chromosomes 2AS, 2AL, 5BS, 6AL, and 6BL. Two of the loci are likely *Lr52* and *Lr64*, while the remaining four loci are most likely novel. Except *QLrdu.2AS*, the identified leaf rust resistance loci were race

specific. For stripe rust, four loci were detected on chromosome arms 1BS, 5BL, and 7BL. All of these loci did not correspond to cataloged *Yr* genes. The loci *QYrdu.5BL-1* and *QYrdu.7BL* were associated with response to the three U.S. *Pst*-races used in this study. For stem rust, 22 resistance loci were detected on chromosomes 1BL, 2BL, 3AL, 3BL, 4AL, 5AL, 5BL, 6AS, 6AL, 6BL, and 7BL. Seven of these *Sr* loci had large effect and high frequencies in this germplasm, thus important to keep in future released durum wheat varieties. Our results showed the presence of known *Sr* genes *Sr8155B1*, *Sr13*, and *Sr7b* that were found together in 40% of this durum wheat germplasm. Seventeen *Sr* loci from this study are not yet cataloged and need to be validated and further characterized. Five of the identified stem rust resistance loci (*QSrdu.1BL*, *QSrdu.2BL*, *QSrdu.5AL-1*, *QSrdu.6AS-1*, and *QSrdu.6BL-3*) were associated with response to more than one race. The novel resistance loci identified in this study will enhance breeding for rust resistance in durum wheat. Because it is relatively easy to make crosses between tetraploid wheat and hexaploid wheat, new rust resistance genes identified in this durum wheat germplasm could also be transferred to common wheat. The SNP markers associated with the large-effect all-stage rust resistance genes/loci in this study can be converted to KASP or STARP markers for use in marker assisted breeding. The presence of gene pyramiding that is already present in this germplasm would be very valuable for breeding for rust resistance.

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

MA and EE conceived and designed the experiments. MA, MR, and JK conducted the experiments. MA and AK analyzed the data. EE provided the resources. MA wrote the manuscript. All authors revised the manuscript.

FUNDING

This work was funded by the North Dakota Wheat Commission, ND, United States.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.640739/full#supplementary-material>

REFERENCES

- Aoun, M., Breiland, M., Turner, K. M., Loladze, A., Chao, S., Xu, S. S., et al. (2016). Genome-wide association mapping of leaf rust response in a durum wheat worldwide germplasm collection. *Plant Genome* 9, 1–24. doi: 10.9734/ajob/2017/38120
- Aoun, M., Kolmer, J. A., Breiland, M., Richards, J., Brueggeman, R. S., Szabo, L. J., et al. (2020). Genotyping-by-sequencing for the study of genetic diversity in *Puccinia triticina*. *Plant Dis.* 104, 752–760. doi: 10.1094/pdis-09-19-1890-re
- Aoun, M., Kolmer, J. A., Rouse, M. N., Chao, S., Bulbula, W. D., Elias, E. M., et al. (2017). Inheritance and bulked segregant analysis of leaf rust and stem rust resistance in durum wheat genotypes. *Phytopathology* 107, 1496–1506. doi: 10.1094/phyto-12-16-0444-r
- Aoun, M., Kolmer, J. A., Rouse, M. N., Elias, E. M., Breiland, M., Bulbula, W. D., et al. (2019). Mapping of novel leaf rust and stem rust resistance genes in the Portuguese durum wheat landrace PI 192051. *G3 Genes Genomes Genet.* 9, 2535–2547. doi: 10.1534/g3.119.400292
- Bansal, U. K., Bossolini, E., Miah, H., Keller, B., Park, R. F., and Bariana, H. S. (2008). Genetic mapping of seedling and adult plant stem rust resistance in two European winter wheat cultivars. *Euphytica* 164, 821–828. doi: 10.1007/s10681-008-9736-z
- Benjamini, Y., and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188.
- Chao, S., Rouse, M. N., Acevedo, M., Szabo-Hever, A., Bockelman, H., Bonman, J. M., et al. (2017). Evaluation of genetic diversity and host resistance to stem rust in USDA NSGC durum wheat accessions. *Plant Genome* 10, 1–13.
- Chen, X. (2005). Epidemiology and control of stripe rust (*Puccinia striiformis* f. sp. *tritici*) on wheat. *Can. J. Plant Pathol.* 27, 314–337.
- Chen, X. (2013). High-temperature adult-plant resistance, key for sustainable control of stripe rust. *Am. J. Plant Sci.* 4, 608–627. doi: 10.4236/ajps.2013.43080
- Cheng, P., Xu, L. S., Wang, M. N., See, D. R., and Chen, X. M. (2014). Molecular mapping of genes *Yr64* and *Yr65* for stripe rust resistance in hexaploid derivatives of durum wheat accessions PI 331260 and PI 480016. *Theor. Appl. Genet.* 127, 2267–2277. doi: 10.1007/s00122-014-2378-8
- Desiderio, F., Guerra, D., Mastrangelo, A. M., Rubiales, D., Pasquini, M., Simeone, R., et al. (2014). “Genetic basis of resistance to leaf rust in tetraploid wheats,” in *Proceedings of the International Symposium on Genetics and Breeding of Durum Wheat*, Vol. 110, eds E. Porceddu, A. B. Damania, and C. O. Qualset (Bari: CIHEAM), 447–452.
- Dyck, P. L. (1994). The transfer of leaf rust resistance from *Triticum turgidum* ssp. *dicoccoides* to hexaploid wheat. *Can. J. Plant Sci.* 74, 671–673. doi: 10.4141/cjps94-121
- Gessese, M., Bariana, H., Wong, D., Hayden, M., and Bansal, U. (2019). Molecular mapping of stripe rust resistance gene *Yr81* in a common wheat landrace Aus27430. *Plant Dis.* 103, 1166–1171. doi: 10.1094/pdis-06-18-1055-re
- Ghosh, J. K., Delampady, M., and Samanta, T. (2006). *An Introduction to Bayesian Analysis: Theory and Methods*. New York, NY: Springer-Verlag.
- Goyeau, H., Berder, J., Czerepak, C., Gautier, A., Lanen, C., and Lannou, C. (2012). Low diversity and fast evolution in the population of *Puccinia triticina* causing durum wheat leaf rust in France from 1999 to 2009, as revealed by an adapted differential set. *Plant Pathol.* 61, 761–772. doi: 10.1111/j.1365-3059.2011.02554.x
- Herrera-Foessel, S., Singh, R. P., Huerta-Espino, J., Crossa, J., Yuen, J., and Djurle, A. (2006). Effect of leaf rust on grain yield and yield traits of durum wheats with race-specific and slow-rusting resistance to leaf rust. *Plant Dis.* 90, 1065–1072. doi: 10.1094/pd-90-1065
- Herrera-Foessel, S. A., Huerta-Espino, J., Calvo-Salazar, V., Lan, C. X., and Singh, R. P. (2014). *Lr72* confers resistance to leaf rust in durum wheat cultivar Atil C2000. *Plant Dis.* 98, 631–635. doi: 10.1094/pdis-07-13-0741-re
- Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., William, H. M., Djurle, A., and Yuen, J. (2008a). Molecular mapping of a leaf rust resistance gene on the short arm of chromosome 6B of durum wheat. *Plant Dis.* 92, 1650–1654. doi: 10.1094/pdis-92-12-1650
- Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., William, H. M., Garcia, V., Djurle, A., et al. (2008b). Identification and molecular characterization of leaf rust resistance gene *Lr14a* in durum wheat. *Plant Dis.* 92, 469–473. doi: 10.1094/pdis-92-3-0469
- Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., William, M., Rosewarne, G., Djurle, A., et al. (2007). Identification and mapping of *Lr3* and a linked leaf rust resistance gene in durum wheat. *Crop Sci.* 47, 1459–1466. doi: 10.2135/cropsci2006.10.0663
- Huerta-Espino, J., and Roelfs, A. (1992). Leaf rust on durum wheats. *Vortr. Pflanzenzuchtg.* 24, 100–102.
- Huerta-Espino, J., Singh, R. P., Herrera-Foessel, S. A., Perez-Lopez, J. B., and Figueroa-Lopez, P. (2009). First detection of virulence in *Puccinia triticina* to resistance genes *Lr27+* *Lr31* present in durum wheat in Mexico. *Plant Dis.* 93, 110–110. doi: 10.1094/pdis-93-1-0110c
- Hundie, B., Girma, B., Tadesse, Z., Edae, E., Olivera, P., Abera, E. H., et al. (2019). Characterization of Ethiopian wheat germplasm for resistance to four *Puccinia graminis* f. sp. *tritici* races facilitated by single-race nurseries. *Plant Dis.* 103, 2359–2366. doi: 10.1094/pdis-07-18-1243-re
- Jin, Y. (2005). Races of *Puccinia graminis* identified in United States during 2003. *Plant Dis.* 89, 1125–1127. doi: 10.1094/pd-89-1125
- Jin, Y., and Singh, R. P. (2006). Resistance in U.S. wheat to recent Eastern African isolates of *Puccinia graminis* f. sp. *tritici* with virulence to resistance gene *Sr31*. *Plant Dis.* 90, 476–480. doi: 10.1094/pd-90-0476
- Jin, Y., Singh, R. P., Ward, R. W., Wanyera, R., Kinyua, M., Njau, P., et al. (2007). Characterization of seedling infection types and adult plant infection responses of monogenic *Sr* gene lines to race TTKS of *Puccinia graminis* f. sp. *tritici*. *Plant Dis.* 91, 1096–1099. doi: 10.1094/pdis-91-9-1096
- Jin, Y., Szabo, L. J., Pretorius, Z. A., Singh, R. P., Ward, R., and Fetch, T. Jr. (2008). Detection of virulence to resistance gene *Sr24* within race TTKS of *Puccinia graminis* f. sp. *tritici*. *Plant Dis.* 92, 923–926. doi: 10.1094/pdis-92-6-0923
- Johnson, M., Kumar, A., Oladzadabbasabadi, A., Salsman, E., Aoun, M., Manthey, F., et al. (2019). Association mapping for 24 traits related to protein content, gluten strength, color, cooking and milling quality using balanced and unbalanced data in durum wheat [*Triticum turgidum* L. var. *Durum* (Desf.)]. *Front. Genet.* 10:1717. doi: 10.3389/fgene.2019.00717
- Klindworth, D. L., Miller, J. D., Jin, Y., and Xu, S. S. (2007). Chromosomal locations of genes for stem rust resistance in monogenic lines derived from tetraploid wheat accession ST464. *Crop Sci.* 47, 1441–1450. doi: 10.2135/cropsci2006.05.0345
- Knott, D. R. (1962). The inheritance of rust resistance: IX. The inheritance of resistance to races 15B and 56 of stem rust in the wheat variety Khapstein. *Can. J. Plant Sci.* 42, 415–419. doi: 10.4141/cjps62-068
- Knott, D. R. (1990). Near-isogenic lines of wheat carrying genes for stem rust resistance. *Crop Sci.* 30, 901–905. doi: 10.2135/cropsci1990.0011183x003000040029x
- Kolmer, J. (2013). Leaf rust of wheat: pathogen biology, variation and host resistance. *Forests* 4, 70–84. doi: 10.3390/f4010070
- Kolmer, J. A. (2015a). First report of a wheat leaf rust (*Puccinia triticina*) phenotype with high virulence to durum wheat in the great plains region of the United States. *Plant Dis.* 99:156. doi: 10.1094/pdis-06-14-0667-pdn
- Kolmer, J. A. (2015b). Leaf rust resistance in wheat line RL6062 is an allele at the *Lr3* locus. *Crop Sci.* 55, 2186–2190. doi: 10.2135/cropsci2015.01.0031
- Kolmer, J. A., and Acevedo, M. (2016). Genetically divergent types of the wheat leaf fungus *Puccinia triticina* in Ethiopia, a center of tetraploid wheat diversity. *Phytopathology* 106, 380–385. doi: 10.1094/phyto-10-15-0247-r
- Kolmer, J. A., Bernardo, A., Bai, G., Hayden, M. J., and Anderson, J. A. (2019). Thatcher wheat line RL6149 carries *Lr64* and a second leaf rust resistance gene on chromosome 1DS. *Theor. Appl. Genet.* 132, 2809–2814. doi: 10.1007/s00122-019-03389-y
- Kolmer, J. A., Herman, A., Ordoñez, M. E., German, S., Morgounov, A., Pretorius, Z., et al. (2020). Endemic and panglobal genetic groups, and divergence of host-associated forms in worldwide collections of the wheat leaf rust fungus *Puccinia triticina* as determined by genotyping by sequencing. *Heredity* 124, 397–409. doi: 10.1038/s41437-019-0288-x
- Kolmer, J. A., and Hughes, M. E. (2013). Physiologic specialization of *Puccinia triticina* on wheat in the United States in 2011. *Plant Dis.* 97, 1103–1108. doi: 10.1094/PDIS-11-12-1068-SR
- Kolmer, J. A., and Hughes, M. E. (2014). Physiologic specialization of *Puccinia triticina* on wheat in the United States in 2012. *Plant Dis.* 98, 1145–1150. doi: 10.1094/pdis-12-13-1267-sr

- Kthiri, D., Loladze, A., MacLachlan, P. R., N'Diaye, A., Walkowiak, S., Nilsen, K., et al. (2018). Characterization and mapping of leaf rust resistance in four durum wheat cultivars. *PLoS One* 13:e0197317. doi: 10.1371/journal.pone.0197317
- Kthiri, D., Loladze, A., N'Diaye, A., Nilsen, K. T., Walkowiak, S., Dreisigacker, S., et al. (2019). Mapping of genetic loci conferring resistance to leaf rust from three globally resistant durum wheat sources. *Front. Plant Sci.* 10:1247. doi: 10.3389/fpls.2019.01247
- Letta, T., Maccaferri, M., Badebo, A., Ammar, K., Ricci, A., Crossa, J., et al. (2013). Searching for novel sources of field resistance to Ug99 and Ethiopian stem rust races in durum wheat via association mapping. *Theor. Appl. Genet.* 126, 1237–1256. doi: 10.1007/s00122-013-2050-8
- Letta, T., Olivera, P., Maccaferri, M., Jin, Y., Ammar, K., Badebo, A., et al. (2014). Association mapping reveals novel stem rust resistance loci in durum wheat at the seedling stage. *Plant Genome* 7, 1–13.
- Lin, X., N'Diaye, A., Walkowiak, S., Nilsen, K. T., Cory, A. T., Haile, J., et al. (2018). Genetic analysis of resistance to stripe rust in durum wheat (*Triticum turgidum* L. var. *durum*). *PLoS One* 13:e0203283. doi: 10.1371/journal.pone.0203283
- Line, R. F., and Qayoum, A. (1992). *Virulence, Aggressiveness, Evolution, and Distribution of Races of Puccinia striiformis (The Cause of Stripe Rust of Wheat) in North America, 1968–87*. Technical Bulletin No. 1788. Washington, DC: United States Department of Agriculture.
- Liu, W., Maccaferri, M., Bulli, P., Rynearson, S., Tuberosa, R., Chen, X., et al. (2017). Genome-wide association mapping for seedling and field resistance to *Puccinia striiformis* f. sp. *tritici* in elite durum wheat. *Theor. Appl. Genet.* 130, 649–667. doi: 10.1007/s00122-016-2841-9
- Loladze, A., Kthiri, D., Pozniak, C., and Ammar, K. (2014). Genetic analysis of leaf rust resistance in six durum wheat genotypes. *Phytopathology* 104, 1322–1328. doi: 10.1094/phyto-03-14-0065-r
- Long, D. L., and Kolmer, J. A. (1989). A North American system of nomenclature for *Puccinia recondita* f.sp. *tritici*. *Phytopathology* 79, 525–529.
- Ma, J. X., Zhou, R. H., Dong, Y. S., Wang, L. F., Wang, X. M., and Jia, J. Z. (2001). Molecular mapping and detection of the yellow rust resistance gene *Yr26* in wheat transferred from *Triticum turgidum* L. using microsatellite markers. *Euphytica* 120, 219–226.
- Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895.
- Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288
- Maccaferri, M., Sanguineti, M. C., Mantovani, P., Demontis, A., Massi, A., Ammar, K., et al. (2010). Association mapping of leaf rust response in durum wheat. *Mol. Breed.* 26, 189–228. doi: 10.1007/s11032-009-9353-0
- Macer, R. C. F. (1966). “The formal and monosomic genetic analysis of stripe rust (*Puccinia striiformis*) resistance in wheat,” in *Proceedings of the 2nd International Wheat Genetics Symposium*. (Suppl. 2) 19–24 August 1963, ed. J. MacKey (Lund: Hereditas), 127–142.
- Magallanes-López, A. M., Ammar, K., Morales-Dorantes, A., González-Santoyo, H., Crossa, J., and Guzmán, C. (2017). Grain quality traits of commercial durum wheat varieties and their relationships with drought stress and glutenins composition. *J. Cereal Sci.* 75, 1–9. doi: 10.1016/j.jcs.2017.03.005
- Marais, G. F., Pretorius, Z. A., Wellings, C. R., McCallum, B., and Marais, A. S. (2005). Leaf rust and stripe rust resistance genes transferred to common wheat from *Triticum dicoccoides*. *Euphytica* 143, 115–123. doi: 10.1007/s10681-005-2911-6
- McFadden, E. S. (1939). Brown necrosis, a discoloration associated with rust infection in certain rust resistance wheats. *J. Agric. Res.* 58, 805–819.
- McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., et al. (2009). *Catalogue of Gene Symbols for Wheat. 2009 Supplement. KOMUGI Wheat Genetic Resources Database, Yokohama, Japan*. Available online at: <https://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp> (accessed September 8, 2019).
- McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., and Xia, X. C. (2017). *Catalogue of Gene Symbols for Wheat: 2017 Supplement*. Available online at: <https://shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement2017.pdf> (accessed 25 June 2019).
- McIntosh, R. A., and Lagudah, E. S. (2000). Cytogenetical studies in wheat. XVIII. Gene *Yr24* for resistance to stripe rust. *Plant Breed.* 119, 81–93. doi: 10.1046/j.1439-0523.2000.00449.x
- McIntosh, R. A., Wellings, C. R., and Park, R. F. (1995). *Wheat Rusts: An Atlas of Resistance Genes*. Melbourne, Vic: CSIRO Publishing, 199.
- McIntosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, J., Morris, C., Appels, R., et al. (2013). *Catalogue of Gene Symbols for Wheat*. Available online at: <https://shigen.nig.ac.jp/wheat/komugi/genes/macgene/2013/GeneSymbol.pdf> (accessed July 20, 2018).
- Mengistu, D. K., and Pè, M. E. (2016). Revisiting the ignored Ethiopian durum wheat (*Triticum turgidum* var. *durum*) landraces for genetic diversity exploitation in future wheat breeding programs. *J. Plant Breed. Crop Sci.* 8, 45–59. doi: 10.5897/jpbcs2015.0542
- Milus, E. A., Kristensen, K., and Hvömmøller, M. S. (2009). Evidence for increased aggressiveness in a recent widespread strain of *Puccinia striiformis* f.sp. *tritici* causing stripe rust of wheat. *Phytopathology* 99, 89–94. doi: 10.1094/phyto-99-1-0089
- Mishra, A. N., Kaushal, K., Dubey, V. G., and Prasad, S. V. (2015). Diverse sources of resistance to leaf rust in durum wheat. *Indian J. Genet. Plant Breed.* 75, 336–340. doi: 10.5958/0975-6906.2015.00053.x
- NASS (2018). *National Agricultural Statistics Service*. Available online at: <https://www.nass.usda.gov> (accessed September 8, 2018).
- Newcomb, M., Olivera, P. D., Rouse, M. N., Szabo, L. J., Johnson, J., Gale, S., et al. (2016). Kenyan isolates of *Puccinia graminis* f. sp. *tritici* from 2008 to 2014: virulence to *SrTmp* in the Ug99 race group and implications for breeding programs. *Phytopathology* 106, 729–736. doi: 10.1094/phyto-12-15-0337-r
- Nirmala, J., Saini, J., Newcomb, M., Olivera, P., Gale, S., Klindworth, D., et al. (2017). Discovery of a novel stem rust resistance allele in durum wheat that exhibits differential reactions to Ug99 isolates. *G3 Genes Genomes Genet.* 7, 3481–3490. doi: 10.1534/g3.117.300209
- Olivera, P., Newcomb, M., Szabo, L. J., Rouse, M., Johnson, J., Gale, S., et al. (2015). Phenotypic and genotypic characterization of race TKTF of *Puccinia graminis* f. sp. *tritici* that caused a wheat stem rust epidemic in southern Ethiopia in 2013–14. *Phytopathology* 105, 917–928. doi: 10.1094/phyto-11-14-0302-fi
- Olivera, P. D., Jin, Y., Rouse, M., Badebo, A., Fetch, T. Jr., Singh, R. P., et al. (2012). Races of *Puccinia graminis* f. sp. *tritici* with combined virulence to *Sr13* and *Sr9e* in a field stem rust screening nursery in Ethiopia. *Plant Dis.* 96, 623–628. doi: 10.1094/pdis-09-11-0793
- Ordoñez, M. E., and Kolmer, J. A. (2007a). Virulence phenotypes of a worldwide collection of *Puccinia triticina* from durum wheat. *Phytopathology* 97, 344–351. doi: 10.1094/phyto-97-3-0344
- Ordoñez, M. E., and Kolmer, J. A. (2007b). Simple sequence repeat diversity of a worldwide collection of *Puccinia triticina* from durum wheat. *Phytopathology* 97, 574–583. doi: 10.1094/phyto-97-5-0574
- Peng, J. H., Fahima, T., Röder, M. S., Huang, Q. Y., Dahan, A., Li, Y. C., et al. (2000). High-density molecular map of chromosome region harboring stripe-rust resistance genes *YrH52* and *Yr15* derived from wild emmer wheat, *Triticum dicoccoides*. *Genetica* 109, 199–210.
- Qureshi, N., Bariana, H., Kumran, V. V., Muruga, S., Forrest, K. L., Hayden, M. J., et al. (2018). A new leaf rust resistance gene *Lr79* mapped in chromosome 3BL from the durum wheat landrace Aus26582. *Theor. Appl. Genet.* 131, 1091–1098. doi: 10.1007/s00122-018-3060-3
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rosewarne, G. M., Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., Lan, C. X., and He, Z. H. (2013). Quantitative trait loci of stripe rust resistance in wheat. *Theor. Appl. Genet.* 126, 2427–2449.
- Rouse, M. N., Nirmala, J., Jin, Y., Chao, S., Fetch, T. G., Pretorius, Z. A., et al. (2014). Characterization of *Sr9h*, a wheat stem rust resistance allele effective to Ug99. *Theor. Appl. Genet.* 127, 1681–1688. doi: 10.1007/s00122-014-2330-y
- Rouse, M. N., Wanyera, R., Njau, P., and Jin, Y. (2011). Sources of resistance to stem rust race Ug99 in spring wheat germplasm. *Plant Dis.* 95, 762–766. doi: 10.1094/pdis-12-10-0940
- Royo, C., Elias, E., and Manthey, F. (2009). “Durum wheat breeding,” in *Cereals. Handbook of Plant Breeding*, Vol. 3, ed. M. Carona (New York, NY: Springer), doi: 10.1007/978-0-387-72297-9_6

- Saini, J., Faris, J. D., Zhang, Q., Rouse, M. N., Jin, Y., Long, Y., et al. (2018). Identification, mapping, and marker development of stem rust resistance genes in durum wheat 'Lebsock'. *Mol. Breed.* 38:77.
- SAS Institute Inc (2004). *SAS/ETS 9.1 User's Guide*. Cary, NC: SAS Institute Inc.
- Singh, B., Bansal, U. K., Forrest, K. L., Hayden, M. J., Hare, R. A., and Bariana, H. S. (2010). Inheritance and chromosome location of leaf rust resistance in durum wheat cultivar Wollaroi. *Euphytica* 175, 351–355. doi: 10.1007/s10681-010-0179-y
- Singh, R. (1991). Pathogenic variations of *Puccinia recondita* f. sp. *tritici* in wheat-growing areas of Mexico during 1988 and 1989. *Plant Dis.* 75, 790–794.
- Singh, R. P., Hodson, D. P., Huerta-Espino, J., Jin, Y., Bhavani, S., Njau, P., et al. (2011). The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. *Annu. Rev. Phytopathol.* 49, 465–481. doi: 10.1146/annurev-phyto-072910-095423
- Singh, R. P., Hodson, D. P., Jin, Y., Lagudah, E. S., Ayliffe, M. A., Bhavani, S., et al. (2015). Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. *Phytopathology* 105, 872–884. doi: 10.1094/phyto-01-15-0030-fi
- Singh, R. P., Huerta-Espino, J., Pfeiffer, W., and Figueroa-Lopez, P. (2004). Occurrence and impact of a new leaf rust race on durum wheat in northwestern Mexico from 2001 to 2003. *Plant Dis.* 88, 703–708. doi: 10.1094/pdis.2004.88.7.703
- Solh, M., Nazari, K., Tadesse, W., and Wellings, C. R. (2012). "The growing threat of stripe rust worldwide," in *Proceedings of the Borlaug Global Rust Initiative (BGRI) Conference*, Beijing, 1–4.
- Stakman, E. C., Stewart, D. M., and Loegering, W. Q. (1962). *Identification of Physiologic Races of Puccinia graminis* var. *tritici*. US Department of Agriculture Agricultural Research Service E-617. Washington DC: United States Department of Agriculture.
- Uauy, C., Brevis, J. C., Chen, X., Khan, I., Jackson, L., Chicaiza, O., et al. (2005). High-temperature adult-plant (HTAP) stripe rust resistance gene *Yr36* from *Triticum turgidum* ssp. *dicoccoides* is closely linked to the grain protein content locus *Gpc-B1*. *Theor. Appl. Genet.* 112, 97–105. doi: 10.1007/s00122-005-0109-x
- USDA, NASS, North Dakota Field Office (2019). Available online at: <https://ndwheat.com/uploads/10/whtvr19.pdf> (accessed June 20, 2020).
- Wan, A., Chen, X., and Yuen, J. (2016). Races of *Puccinia striiformis* f. sp. *tritici* in the United States in 2011 and 2012 and comparison with races in 2010. *Plant Dis.* 100, 966–975. doi: 10.1094/pdis-10-15-1122-re
- Wan, A. M., and Chen, X. M. (2014). Virulence characterization of *Puccinia striiformis* f. sp. *tritici* using a new set of *Yr* single-gene line differentials in the United States in 2010. *Plant Dis.* 98, 1534–1542. doi: 10.1094/pdis-01-14-0071-re
- Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183
- Wei, T., and Simko, V. (2013). *corrplot: Visualization of a Correlation Matrix*. R Package Version 0.73 230:11.
- Xu, L. S., Wang, M. N., Cheng, P., Kang, Z. S., Hulbert, S. H., and Chen, X. M. (2013). Molecular mapping of *Yr53*, a new gene for stripe rust resistance in durum wheat accession PI 480148 and its transfer to common wheat. *Theor. Appl. Genet.* 126, 523–533. doi: 10.1007/s00122-012-1998-0
- Zhang, D., Bowden, R. L., Yu, J., Carver, B. F., and Bai, G. (2014). Association analysis of stem rust resistance in US Winter Wheat. *PLoS One* 9:e103747. doi: 10.1371/journal.pone.0103747
- Zhang, W., Chen, S., Abate, Z., Nirmala, J., Rouse, M. N., and Dubcovsky, J. (2017). Identification and characterization of *Sr13*, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. *Proc. Natl. Acad. Sci. U.S.A.* 114, E9483–E9492.
- Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546

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 Aoun, Rouse, Kolmer, Kumar and Elias. 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 Improvement for Resistance to Black Sigatoka in Bananas: A Systematic Review

Julianna M. S. Soares^{1†}, Anelita J. Rocha^{1†}, Fernanda S. Nascimento¹, Adriadna S. Santos², Robert N. G. Miller³, Cláudia F. Ferreira⁴, Fernando Haddad⁴, Vanusia B. O. Amorim⁴ and Edson P. Amorim^{4*}

¹ Department of Biological Sciences, Feira de Santana State University, Feira de Santana, Brazil, ² Secretariat of Education of the State of Bahia, Salvador, Brazil, ³ Department of Cell Biology, University of Brasília, Brasília, Brazil, ⁴ Embrapa Mandioca e Fruticultura, Cruz das Almas, Brazil

OPEN ACCESS

Edited by:

Anna Maria Mastrangelo,
Council for Agricultural and
Economics Research (CREA), Italy

Reviewed by:

Fernando Martinez,
Sevilla University, Spain
Karl Kunert,
University of Pretoria, South Africa

*Correspondence:

Edson P. Amorim
edson.amorim@embrapa.br

[†]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: 24 January 2021

Accepted: 19 March 2021

Published: 21 April 2021

Citation:

Soares JMS, Rocha AJ,
Nascimento FS, Santos AS,
Miller RNG, Ferreira CF, Haddad F,
Amorim VBO and Amorim EP (2021)
Genetic Improvement for Resistance
to Black Sigatoka in Bananas: A
Systematic Review.
Front. Plant Sci. 12:657916.
doi: 10.3389/fpls.2021.657916

Bananas are an important staple food crop in tropical and subtropical regions in Asia, sub-Saharan Africa, and Central and South America. The plant is affected by numerous diseases, with the fungal leaf disease black Sigatoka, caused by *Mycosphaerella fijiensis* Morelet [anamorph: *Pseudocercospora fijiensis* (Morelet) Deighton], considered one of the most economically important phytosanitary problem. Although the development of resistant cultivars is recognized as most effective method for long term control of the disease, the majority of today's cultivars are susceptible. In order to gain insights into this pathosystem, this first systematic literature review on the topic is presented. Utilizing six databases (PubMed Central, Web of Science, Google Academic, Springer, CAPES and Scopus Journals) searches were performed using pre-established inclusion and exclusion criteria. From a total of 3,070 published studies examined, 24 were relevant with regard to the *Musa-P. fijiensis* pathosystem. Relevant papers highlighted that resistant and susceptible cultivars clearly respond differently to infection by this pathogen. *M. acuminata* wild diploids such as Calcutta 4 and other diploid cultivars can harbor sources of resistance genes, serving as parentals for the generation of improved diploids and subsequent gene introgression in new cultivars. From the sequenced reference genome of *Musa acuminata*, although the function of many genes in the genome still require validation, on the basis of transcriptome, proteome and biochemical data, numerous candidate genes and molecules have been identified for further evaluation through genetic transformation and gene editing approaches. Genes identified in the resistance response have included those associated with jasmonic acid and ethylene signaling, transcription factors, phenylpropanoid pathways, antioxidants and pathogenesis-related proteins. Papers in this study also revealed gene-derived markers in *Musa* applicable for downstream application in marker assisted selection. The information gathered in this review furthers understanding of the immune response in *Musa* to the pathogen *P. fijiensis* and is relevant for genetic improvement programs for bananas and plantains for control of black Sigatoka.

Keywords: black Sigatoka, *Musa* spp., *Pseudocercospora fijiensis*, genetic resistance, state-of-the-art

INTRODUCTION

Bananas and plantains (*Musa* spp.) are important commodity fruit crops in terms of trade and consumption, and represent the fourth most important staple food worldwide (Weber et al., 2017). World production in 2018 was ~154.5 million tons, of which 74% were bananas and 26% plantains, grown over a total area of 11.3 million hectares (FAOSTAT, 2021).

Although bananas originated in Southwest Asia and the Western Pacific region, popularity and economic importance occurred following introduction to Africa, Latin and Central America and the South Pacific (Valmayor, 2001; De Langhe et al., 2009). The vast majority of banana and plantain cultivars originated from hybrids of the two wild diploid species, *Musa acuminata* Colla (genome A) and *M. balbisiana* Colla (genome B). Such crossings resulted in a series of diploids, triploids and tetraploids, with genomic groups classified as AA, AB, AAA, AAB, ABB, AABB, AAAB, and ABBB (Simmonds and Shepherd, 1955).

Banana and plantain production are affected by various pests and diseases, including bacterial wilt (Addy et al., 2016), nematodes (Seenivasan, 2017), Fusarium wilt (Dita et al., 2018; Arinaitwe et al., 2019) and yellow and black Sigatoka diseases (Ferreira et al., 2004; Timm et al., 2016). Black Sigatoka, caused by the fungus *Mycosphaerella fijiensis* Morelet [anamorph: *Pseudocercospora fijiensis* (Morelet) Deighton], can result in considerable negative economic impact, affecting both bananas and plantains across all global growing regions. Whilst chemical control is considered efficient, problems can arise from indiscriminate use, where this approach is detrimental to human health and the environment. Agrochemical-based control is also expensive (Churchill, 2011), with data indicating ~US\$ 1,000/ha spent on disease control annually in large plantations, corresponding to up to 30% of the total production costs (Churchill, 2011; Alakonya et al., 2018). Another important factor to be considered with dependency on agrochemicals is the possible medium- and long-term selection for pathogen strains acquiring resistance to fungicides, potentially reducing effectiveness (Churchill, 2011; Chong, 2016; Friesen, 2016; Rodríguez-García et al., 2016; Oiram-Filho et al., 2019).

Rain splash of asexual conidia and airborne dispersal of sexual ascospores enable effective spread of *P. fijiensis* (Churchill, 2011; Rodríguez-García et al., 2016; Alakonya et al., 2018). The onset of the first symptoms of the disease typically occurs between 7 and 14 days after contamination, depending on local environmental conditions. Following fungal penetration of leaf stomata, colonization of intercellular spaces and subsequent necrotic damage then decrease the photosynthetic capacity of the plant, reducing the quantity and quality of fruits (Churchill, 2011; Alakonya et al., 2018; Cruz-Martín et al., 2018).

Whilst increased understanding of the genetic structure of pathogen populations and their evolution are important components to consider in strategies for *Musa* genetic improvement and management of the disease (Churchill, 2011), the identification at the molecular level of host genes related to resistance to *P. fijiensis* will advance improvement of banana through both assisted selection and genetic engineering

(Mendoza-Rodríguez, 2014). Our understanding of the innate immune system in plants has advanced considerably in recent years, with challenge by pathogen molecules known to activate host receptor proteins for pathogen recognition. In a first layer of the immune response, referred to as pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI), or non-host resistance, host cell surface pattern recognition receptors (PRRs) (Dangl and Jones, 2001; Monaghan and Zipfel, 2012) recognize conserved pathogen-associated molecular patterns (PAMPs) (Jones and Dangl, 2006; Boutrot and Zipfel, 2017) such as bacterial flagellin and fungal cell wall chitin (Felix et al., 1999; Wan et al., 2004; Thomma et al., 2011; Zipfel, 2014; Gong et al., 2020). Plant PRRs, which include receptor-like kinases (RLKs) and receptor-like proteins (RLPs), generally contain extracellular domains with a capacity for ligand binding, transmembrane domains and intracellular domains (Zipfel, 2014). Activation of PRRs following PAMP recognition will trigger intracellular signaling and plant defense responses to block pathogen advance in the host. These include reactive oxygen species (ROS), mitogen-activated protein kinase (MAPK) cascades and Ca^{2+} signaling influx (Chisholm et al., 2006; Dangl et al., 2013; Li et al., 2016). Race-specific pathogen effector proteins, or avirulence (Avr) proteins, when secreted into the host cell by evolving pathogens, by contrast, can suppress PTI and result in an effector-triggered susceptibility (ETS) with subsequent disease (Jones and Dangl, 2006; Boller and Felix, 2009). In a second layer of the plant immunity defense response, intracellular nucleotide-binding and leucine-rich repeat domain intracellular resistance receptors (NLRs) recognize directly or indirectly evolved pathogen effectors, activating effector-triggered immunity (ETI) (Jones and Dangl, 2006). As a more intense response, this again involves calcium ion signaling and ROS, together with transcriptional reprogramming, changes in levels of plant hormones salicylic acid (SA) and jasmonic acid (JA) (Creelman and Mullet, 1995), and the accumulation of pathogenesis-related (PR) proteins (Gururani et al., 2012). Such a suite of responses can also involve the signature hypersensitive response, comprising a programmed and localized host cell death at the site of infection (Jones and Dangl, 2006; Coll et al., 2011; Cui et al., 2015), effectively limiting pathogen advance. Subsequent systemic acquired resistance (SAR) can also occur, conferring a broad spectrum response in the host that heightens resistance to any subsequent pathogen attack (Dong, 2001; Spoel and Dong, 2012).

The pathosystem *Musa* spp. x *P. fijiensis* is complex, given the characteristics of the polyploid host and the morphophysiology of the hemibiotrophic fungus. To date, there have been few studies on the biology of this hemibiotroph and the mode of action of genes involved in the host-pathogen interaction (Cavalcante et al., 2011; Torres et al., 2012; Mendoza-Rodríguez, 2014; Arango-Isaza et al., 2016). Similarly, although the genus *Musa* has been relatively widely studied with regard to molecular marker development and analysis of genetic diversity, with whole genome sequences also developed in recent years for *M. acuminata* and related species, detailed investigation and validation of gene function in immune responses in different *Musa*-pathogen interactions remains limited (Sun et al., 2009;

Li et al., 2012; Wang et al., 2012; Bai et al., 2013; Castañeda et al., 2017). With regard to *Musa-Pseudocercospora* interactions, candidate gene discovery has broadly been undertaken through analysis of gene analogs and through transcriptomics approaches (Miller et al., 2008, 2011; Emediato et al., 2009, 2013; Portal et al., 2011; D' Hont et al., 2012; Passos et al., 2012, 2013; Sulliman et al., 2012; Timm et al., 2016).

Systematic literature reviews are analyses that gather and critically evaluate compiled data from previously published scientific investigations. Such an approach for synthesis of findings is widely employed in medical fields, enabling, in a single document, relevant information to be gathered on a specific topic, for example on a disease or active ingredient in medicines and potential side effects (Falcomer et al., 2019; Jones et al., 2020). For *Musa spp.*, there have only been two studies using such a strategy, with focus on plant physiology associated with water deficit and on fruit consumption preferences (Santos et al., 2018; Falcomer et al., 2019).

Accumulation of knowledge on host genetics and genomics, resistance and defense mechanisms, together with information on methods and tools employed in development of resistance to black Sigatoka, is relevant for genetic improvement strategies for development of resistant cultivars. This systematic review synthesizes relevant literature published in the last 10 years on genetic improvement of banana with a focus on black Sigatoka, to answer the following question: what are the strategies adopted in genetic improvement that aim to reduce the impact of black Sigatoka on banana plants? To our knowledge, this is the first systematic review applied to the *Musa spp.* x *P. fijiensis* pathosystem.

MATERIALS AND METHODS

The systematic review was conducted using the software StArt (State of the Art through Systematic Review) Beta version. 3.0.3, developed at the Federal University of São Carlos (UFSCar) to assist in systematic reviewing (Santos et al., 2018). The software is freely available at http://lapes.dc.ufscar.br/tools/start_tool. This review consisted of three fundamental steps, summarized in **Figure 1**.

Planning

In this step, a defined protocol was followed according to the following information: article title, authors, objective, keywords, research questions, research sources, inclusion/exclusion criteria and definition of study type (<https://doi.org/10.5281/zenodo.4437073>). The questions raised in this review are listed in **Table 1**.

Execution

In order to answer the question of our research, “which strategies were adopted in genetic breeding to reduce the impact of black Sigatoka in bananas?” a research strategy of Population Intervention Comparison Outcome (PICOS), was used (de Costa Santos et al., 2007). This strategy guides what the research question really needs to specify avoiding a less biased answer (Wright et al., 2007). For its elaboration, these following questions should be answered:

P–What is the research problem or who are the individuals populations?

I–What will be done, or which treatment or intervention or exposure?

C–Will any action intervention alternative treatment, or in parallel, be carried out?

O–What is the expected result or outcome?

S–What is the type of study?

The PICOS strategy used in this systematic review is shown in **Table 2**.

Searches were conducted in selected databases: CAPES journals (<https://www.periodicos.capes.gov.br>), PubMed Central (<https://www.ncbi.nlm.nih.gov/pmc>), Google Scholar (<https://scholar.google.com.br>), Springer (<https://link.springer.com>), Web of Science (<http://apps.isiknowledge.com>) and Scopus (<https://www.scopus.com>). The selected files were imported in BIBTEX and MEDLINE format compatible with StArt. Automated searches were made from the themes located in titles, keywords and summaries. Additional articles of relevance that were not identified automatically were subsequently added manually. For all databases, the same search string was employed, with connectors such as “or” and “and” used to group synonymous keywords and the main topics. The String employed was as follows: *Musa spp.* and bananas or plantains and black Sigatoka or *Mycosphaerella fijiensis* or *Pseudocercospora fijiensis* and genetic resistance and markers and genes.

Summarization

This step comprised the elaboration of graphs, tables and a word cloud to summarize the systematic review. All articles that were selected during the selection and extraction phase were based on the following inclusion criteria: articles that contained the search string terms in the title, abstract or keywords; and articles that answered the protocol questions (**Table 1**). Criteria for exclusion were as follows: theses, dissertations, manuals, reports, book chapters, review articles, articles published in annals of events and studies without any clear contribution.

During the selection stage, articles imported into the software StArt were classified as accepted, rejected, or excluded due to duplication. In the extraction phase, a second selection was made considering only the articles that were accepted in the initial selection stage. During this phase, it was possible to delete duplicates, accept articles or reject those that were not in accordance with the objectives of the work, based on reading the articles in full, as well as on the inclusion and exclusion criteria. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist is presented for download at <https://doi.org/10.5281/zenodo.4659141>.

RESULTS

Database Searches

Aiming to reduce bias risks we opted to insert only articles with scientific and statistical data and also those which really considered our main and secondary questions whose conclusions were reliable. Regarding the specific evaluation of risk and bias tools used in clinical studies that were not yet adapted

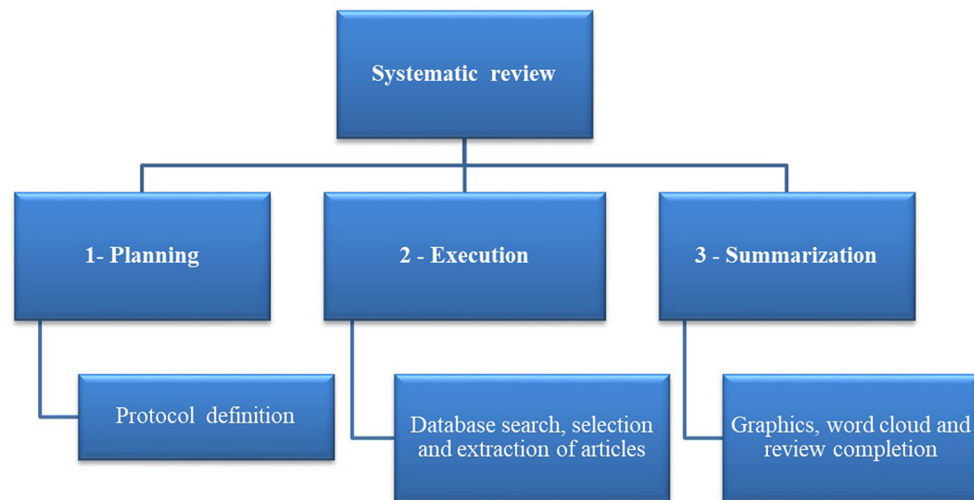


FIGURE 1 | General systematic literature review flowchart [Adapted from Santos et al. (2018)].

TABLE 1 | List of questions raised in the review.

Questions

1. Which countries carried out most studies on the genetic improvement of banana related to black Sigatoka?
2. Which institutions/bodies work with this theme?
3. Which are the most studied *Musa* genotypes and varieties?
4. In terms of commercial cultivars, which are resistant and which are susceptible to *P. fijiensis*?
5. What types of trials are proposed in the studies?
6. Which genes are reported to be associated with resistance to black Sigatoka?
7. What are the biotechnological techniques employed in the studies?
8. What are the structural, genetic and molecular mechanisms involved in *Musa* defense responses responsible for conferring resistance to black Sigatoka?

for use in other areas of knowledge also related to meta-analysis, were used. A PRISMA checklist was also used which is used strategically in systematic reviews aiming transparency and quality in the elaboration and publication of this review. Therefore, we guarantee that there is no bias risks in our review since all the PRISMA parameters were followed accordingly, guaranteeing reproducibility and reliability.

Electronic database searches using StArt resulted in the selection of 3,070 articles, published between January 2010 and December 2020. PubMed Central contributed with the largest number for this systematic review, corresponding to 1,786 papers, or 58% of the total. Web of Science contributed with 1,130 papers, representing 37% of those initially selected, followed by Google Academic (102), Springer (47), CAPES Journal (4) and Scopus (3). Although papers were selected using the search string, most were subsequently excluded from the study, as they were not related to the topic, and/or falling within the exclusion criteria. Two articles were also added manually (Figure 2).

During the initial evaluation of articles based on title and abstract, 2,070 articles did not meet the inclusion criteria. Together with 142 articles that were duplicated, these were all excluded from the systematic review. In the extraction stage, of the 228 remaining articles, 24 were accepted for analysis in

the review from the criteria established for inclusion, as these answered the questions proposed in the initial protocol. For consultation purposes, these are stored in a free digital library at the following link: <https://doi.org/10.5281/zenodo>.

A word cloud was generated during the extraction phase of the database search based on the frequency of keywords in the selected articles ($n = 228$). Highest frequencies of keywords in the articles were observed for black Sigatoka, *Mycosphaerella fijiensis*, *Musa* spp., disease and genetic resistance (Figure 3).

Study Locations

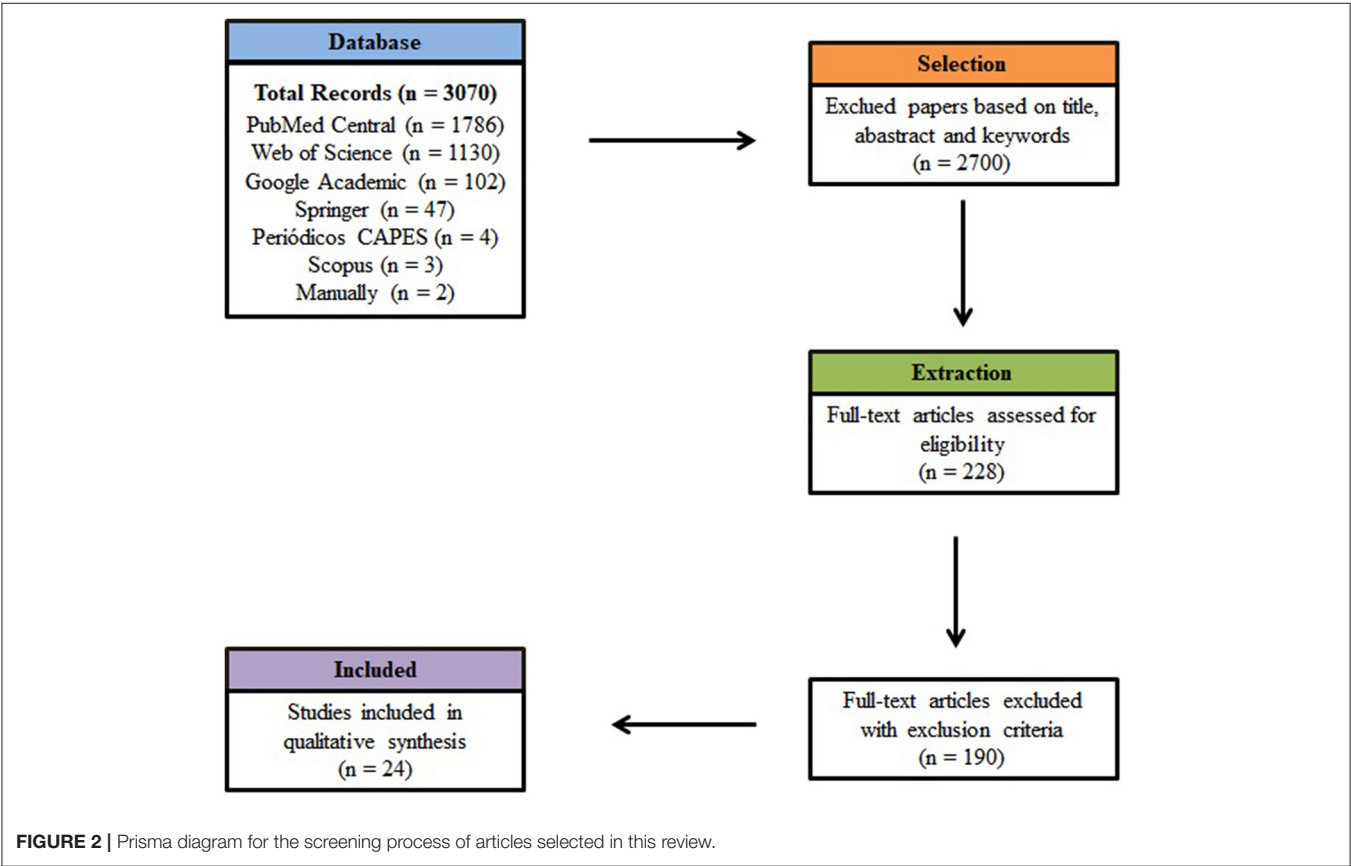
Most of the research work included in this systematic review was conducted in only three countries, namely Cuba (21%), Brazil (18%) and Colombia (17%). Belonging to the American continent, these represented the source of ~67% of the total 24 articles examined (Figure 4A). Articles from Africa, Europe and Asia represented 17, 13, and 4%, respectively (Figure 4B).

Sources of Resistance and Study Environment

Cultivars and genotypes that are resistant, moderately resistant or susceptible to black Sigatoka were the object of study across the selected articles (Table 3). As summarized in Figure 5, most

TABLE 2 | Definition of the PICO terms of strategy for the question in the research used in this research.

| Description | Abbreviation | Components of the question |
|-----------------------|--------------|---|
| Population | P | Banana plants (<i>Musa</i> spp.) with black Sigatoka |
| Interest/intervention | I | Genetic breeding methods used to control the disease. |
| Comparison | C | Lack of breeding methods or any other method of management or control of the disease, which does not involve genetic breeding (cultural, chemical, biological or other methods of control and management of the disease). |
| Outcome | O | Resistance or tolerance to black Sigatoka (basal or complete resistance) |
| Type of study | S | Scientific articles with experimental studies. |



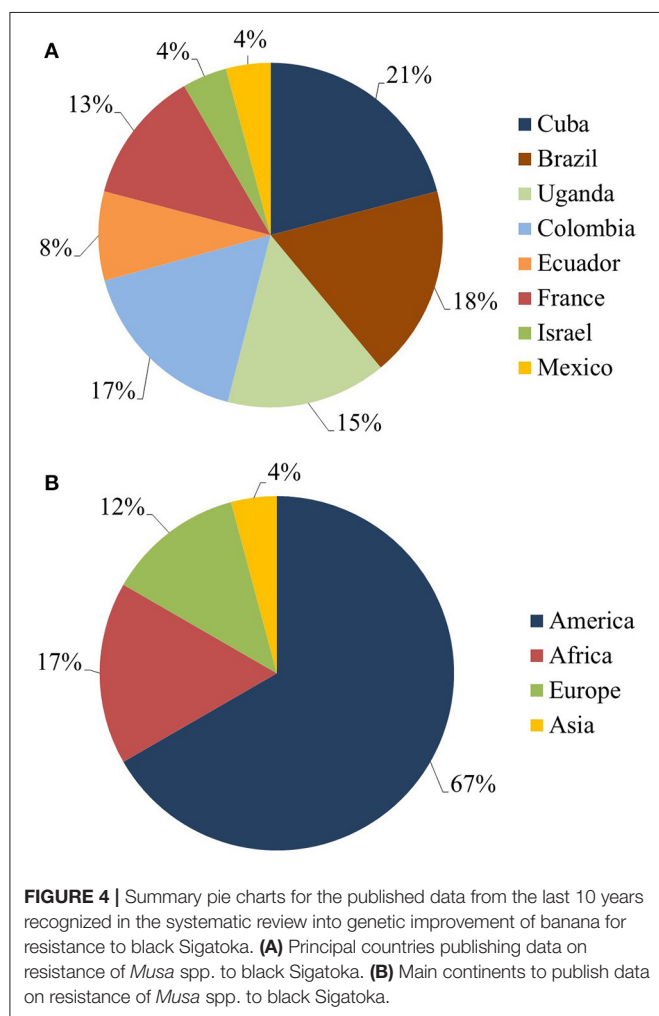
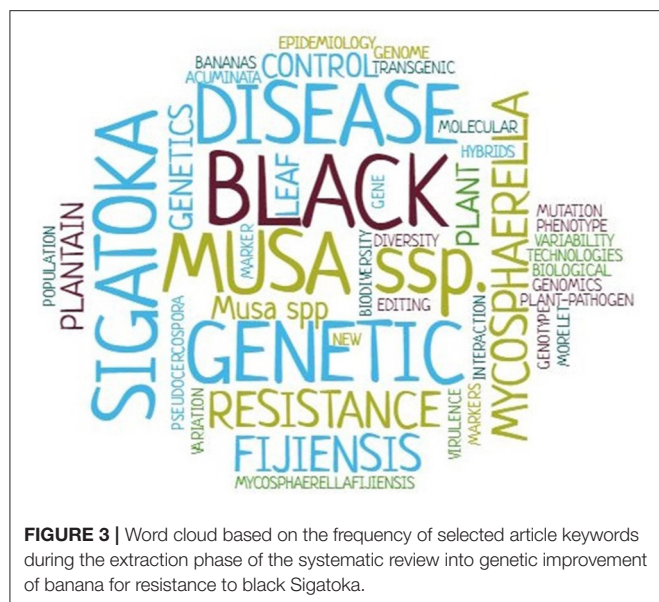
genotypes were diploid AA genome members, representing 46% of those studied, 28% were AAA triploids, 13% AAB genome triploids, 13% AAAB genome tetraploids, and 1% were AB genome diploids. Genotypes most widely employed in studies with *P. fijiensis* were identified as: *M. acuminata* Calcutta 4, Grande Naine and Williams. Although the majority of the resistant or moderately resistant genotypes were AA diploids, resistance was also reported across AAA, AAB, and AAAB members.

With regard to the study environment, most of the studies were conducted on plant material *in vitro* (46%), followed by greenhouse (23%), field-based (15%) and glasshouse environments (13%) (Figure 6). *In vitro* work encompassed laboratory activities such as propagation of plants for testing,

molecular analysis and fungal multiplication. Field work focused on analysis of agronomic characters, evaluation of resistance and other complementary analyses, such as consumer acceptance of resistant cultivars. In relation to greenhouse experiments, different pathogen inoculation approaches were utilized during evaluation of levels of resistance of different banana genotypes to *P. fijiensis* (Alvarado-Capó et al., 2003; Leiva-Mora et al., 2010).

Methodologies Employed

With regard to the main methodologies employed, gene expression analysis was addressed in 38% of the selected publications, followed by enzyme analysis (17%), symptomatology analysis (13%), transgenic development (13%),



agronomic characterization (8%), *Musa* hybridization (8%) and characterization with molecular markers (4%) (Figure 7).

Leaf disease symptom evaluation employed grading scales that were proposed by Alvarado-Capó et al. (2003) and Stover (1972), as modified by Gauhl (1989). Three publications employing transgenic approaches were also identified in the study. Vishnevetsky et al. (2011) focused on the development of a transformation system for banana for pathogen control, with expression of the ThEn - 42 endochitinase gene from *Trichoderma harzianum*, together with a styrene synthase (StSy) gene resulting in transgenic events with improved tolerance to Sigatoka. Onyilo et al. (2018) conducted pathogen gene silencing approaches targeting mitogen-activated protein kinase pathogen genes Fus3 and Slt2, which are reported to be essential for pathogenicity. Portal et al. (2012) verified a green fluorescent protein-transformed *Mycosphaerella fijiensis* strain on susceptible banana “Grande Naine” and resistant “Yangambi km 5” plants, demonstrating that mutation events in *P. fijiensis* can increase virulence. In relation to agronomic characterization, two articles evaluated growth and production performance of genotypes resistant to black Sigatoka (Nowakunda et al., 2015; Weber et al., 2017). Enzymatic activity was also addressed in four publications that reported host enzyme actions during plant-pathogen interaction (Table 4). In two articles, *Musa* interspecific hybridization was also used to assess resistance development to black Sigatoka in progenies (Barekye et al., 2011; Tumuhimbise et al., 2018). Regarding molecular markers, one article addressed the development of microsatellite markers as a resource for *Musa* genetic improvement for resistance (Passos et al., 2012).

Musa Gene Expression Analysis During Interaction With *P. fijiensis*

Overall, eight articles (38%) investigated gene expression during the *Musa* × *P. fijiensis* interaction. Several candidate genes expressed differentially potentially involved in defense responses were identified in the selected articles (Supplementary Table 1). Of the genes identified in this systematic review, 18% are classified as an unassigned function, that is, the functions of these genes have yet to be discovered. The other genes are related to jasmonic acid signaling (14%), ethylene signaling (13%), primary metabolism (8%), secondary metabolism (8%), transcription factors (7%), via phenylpropanoid pathways (6%), antioxidants (6%), carbohydrate metabolism (5%), proteins related to pathogenesis (2%), among others (Figure 8) (Supplementary Table 1). In total, six different methods were used to inoculate the plants, with differences mainly in the form of application of spores on the leaf (brush or spray) and in relation to the concentration of spores, with values ranging from 1×10^3 to 1×10^6 (Supplementary Table 1).

Enzymatic Activity

A total of 10% ($n = 4$) of the articles were related to analysis of enzyme activity in plants infected with *P. fijiensis* (Figure 7). In these publications, increased activity following inoculation was shown for the enzymes

TABLE 3 | *Musa* spp. genotypes most employed in published data recognized in the systematic review into genetic improvement of banana for resistance to black Sigatoka.

| <i>Musa</i> genotype | Genomic group | Classification |
|----------------------|---------------|----------------------|
| Calcutta 4 | AA | Resistant |
| Orito | AA | Resistant |
| Birmanie | AA | Resistant |
| Krasan Saichon | AA | Resistant |
| Tuu Gia | AA | Resistant |
| Zebrina | AA | Resistant |
| N° 118 | AA | Resistant |
| DH-Pahang | AA | Resistant |
| Pisang Lilin | AA | Moderately resistant |
| 028003-01 | AA | Moderately resistant |
| Buitenzorg | AA | Moderately resistant |
| Khi Maeo | AA | Moderately resistant |
| M53 | AA | Moderately resistant |
| Malaccensis 1 | AA | Moderately resistant |
| Malaccensis 2 | AA | Moderately resistant |
| Malbut | AA | Moderately resistant |
| Mambee Thu | AA | Moderately resistant |
| Microcarpa | AA | Moderately resistant |
| Niyama Yik | AA | Moderately resistant |
| PA Rayong | AA | Moderately resistant |
| Pisang Madu | AA | Moderately resistant |
| Pisang Cici | AA | Moderately resistant |
| Pisang Jaran | AA | Moderately resistant |
| Pisang Jari Buaya | AA | Moderately resistant |
| Pisang Lidi | AA | Moderately resistant |
| Pisang Pipit | AA | Moderately resistant |
| Pisang Rojo Uter | AA | Moderately resistant |
| Pisang Tongat | AA | Moderately resistant |
| SF-751 | AA | Moderately resistant |
| Tjau Lagada | AA | Moderately resistant |
| Akondro Mainty | AA | Susceptible |
| Khai Nai On | AA | Susceptible |
| Pisang Berlin | AA | Susceptible |
| Tong Dok Mak | AA | Susceptible |
| IAC 1 | AB | Susceptible |
| Yangambi Km5 | AAA | Resistant |
| Kiwangaazi (M9) | AAA | Resistant |
| Grande naine | AAA | Susceptible |
| Williams | AAA | Susceptible |
| Filipino | AAA | Susceptible |
| Gross Michel | AAA | Susceptible |
| Guineo de seda | AAA | Susceptible |
| Guineo de Jardim | AAA | Susceptible |
| Guineo mulato | AAA | Susceptible |
| Guineo morado | AAA | Susceptible |
| Nakitembe | AAA | Susceptible |
| Limeño | AAB | Resistant |
| NAROBan1 | AAB | Resistant |
| NAROBan2 | AAB | Resistant |

(Continued)

TABLE 3 | Continued

| <i>Musa</i> genotype | Genomic group | Classification |
|----------------------|---------------|----------------|
| NAROBan3 | AAB | Resistant |
| NAROBan4 | AAB | Resistant |
| Thap Maeo | AAB | Resistant |
| Maqueño | AAB | Susceptible |
| Dominico | AAB | Susceptible |
| Dominico gigante | AAB | Susceptible |
| Dominico negro | AAB | Susceptible |
| Dominico-Hartón | AAB | Susceptible |
| Barraganete | AAB | Susceptible |
| PV42-68 | AAAB | Resistant |
| Pacovan Ken | AAAB | Resistant |
| BRS Vitória | AAAB | Resistant |
| BRS Japira | AAAB | Resistant |
| BRS Preciosa | AAAB | Resistant |
| BRS Garantida | AAAB | Resistant |
| BRS Tropical | AAAB | Resistant |
| BRS Platina | AAAB | Resistant |
| BRS Maravilha | AAAB | Resistant |
| FHIA 02 | AAAB | Resistant |
| FHIA 18 | AAAB | Resistant |

peroxidase (POX), phenylalanine ammonia lyase (PAL), β -1, 3-glucanase (GLU) and chitinase (CHI), superoxide dismutase (SOD), ascorbate peroxidase (APX), with elevated H_2O_2 production after infection and pathogen advance also shown (Table 4). In general, enzyme activity was investigated through comparison of resistant and susceptible genotypes after inoculation with *P. fijiensis*. One exception was the publication by Cruz-Martín et al. (2018), where enzymatic activity in *Musa* was analyzed in response to a strain of *Bacillus pumilus*.

DISCUSSION

Database Searches

This review gathered articles published from January 2010 to December 2020 containing information related to studies on the genetic improvement of *Musa* spp. for resistance to *P. fijiensis*. Only articles that answered the questions established in the initial protocol were selected, with emphasis on genetic improvement of bananas and plantains. For this reason, first reports of the disease, articles on the genetic diversity of *P. fijiensis*, and strategies for disease management were not considered in the study. In addition, literature reviews were excluded in order to avoid underestimation of data, as data could theoretically be repeated when considering that the reviews published cite a large number of articles that are already present in our systematic review. In addition, we opted for articles that performed experimental analyses.

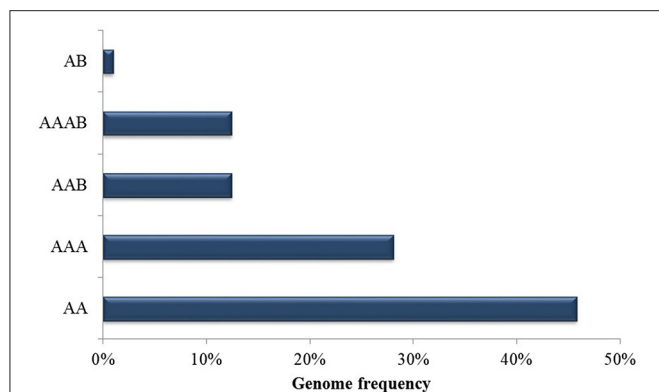


FIGURE 5 | Genotype frequency of *Musa* spp. employed in published data recognized in the systematic review into genetic improvement of banana for resistance to black Sigatoka.

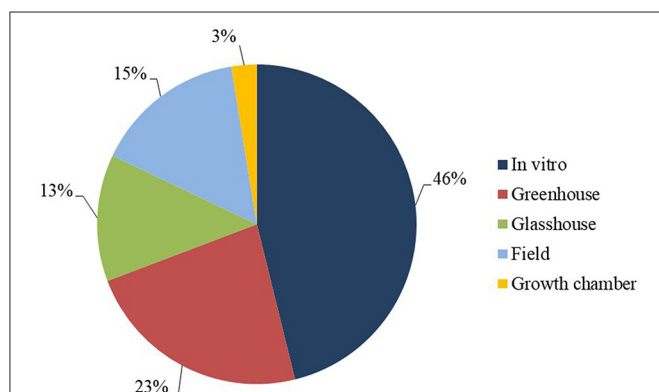


FIGURE 6 | Study environment frequency.

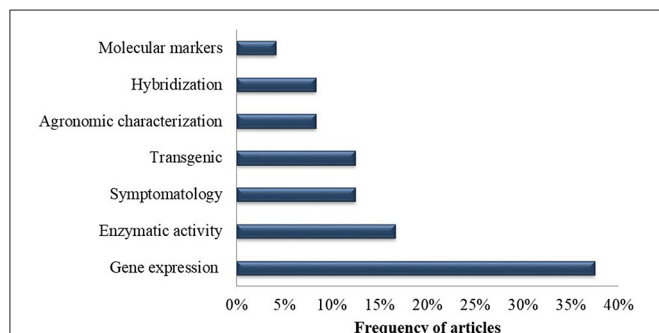


FIGURE 7 | Frequency of methodologies utilized in the selected publications recognized in the systematic review in of genetic improvement of banana for resistance to black Sigatoka.

Study Locations

Latin America accounts for 25% of the world's banana production and 80% of banana exports (FAOSTAT, 2021). Although Brazil is ranked fourth in terms of global banana production, production in the country is destined almost entirely to internal markets.

Brazil and Cuba stand out in this study with the largest number of studies conducted within the objectives of this review. These countries, in addition to having adequate climates for the development of *P. fijiensis*, both employ irrigation systems for banana and plantain cultivation, potentially creating an environment favorable to the fungus. With the exception of certain high-altitude regions (> 1,500 m) (Costa Rica, Guatemala and Mexico), studies have shown that Central America has a natural rain scenario climate which is suitable for the persistence of *P. fijiensis*. In Latin America, Costa Rica is considered the second largest exporter of commercial bananas. Here, however, the black Sigatoka index is high, with fungicides applied up to 45 times a year in heavily infested areas (Yonow et al., 2019). In the main banana export cultivation areas of South America (Northern Colombia, Ecuador and Peru), the climate is less prone to the development of *P. fijiensis* when compared to Central America (Yonow et al., 2019). Here too, however, the number of fungicide cycles has increased considerably, particularly in Ecuador. This is likely due to reduced sensitivity of *P. fijiensis* populations to the widely employed fungicides (Jimenez et al., 2007). A study by Bebbler (2019) on climate change related to black Sigatoka showed that in banana cultivation areas in Latin America and the Caribbean, the risk of infection has increased by a median of 44.2% since 1960. This is likely due to increased humidity and temperatures more favorable to the development of the pathogen. Although increased banana production and global trade have also probably facilitated the establishment and spread of black Sigatoka, climate change has made these regions more conducive to pathogen infection of plants (Bebbler, 2019).

Musa Breeding and Black Sigatoka Resistant Cultivars

The development of black Sigatoka resistant cultivars has been the focus of numerous breeding programs worldwide, with a number employing biotechnology as a support tool. The main banana breeding programs mentioned in the review are located in Africa, Asia and the Americas. In Africa, these comprise the International Institute of Tropical Agriculture (IITA), the National Research Organization (NARO), the Center Africain de Recherches sur Bananiers et Plantains (CARBAP) and the Center National de Recherche Agronomique (CNRA). In Asia, breeding programs are conducted at the National Banana Research Center (NRCB), the Indonesian Fruits Research Institute (ITFRI) and the Chinese Academy of Tropical Agricultural Sciences (CATAS). In the Americas, the Brazilian Agricultural Research Corporation (EMBRAPA), the Honduras Foundation for Agricultural Research (FHIA) and the Center de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) are active in *Musa* improvement. These programs have made significant progress to date in breeding for resistance to black Sigatoka. The FHIA program developed a number of genotypes resistant to *P. fijiensis* that are now grown in different countries around the world, such as Uganda, Tanzania, Ghana, Kenya and Nigeria. In addition, these genotypes have also been employed in breeding programs at IITA, EMBRAPA, CIRAD, and CARBAP (Tenkouano and

TABLE 4 | Enzyme activities in *Musa* spp. during interaction with *Pseudocercospora fijiensis* in the selected publications recognized in the systematic review into genetic improvement of banana for resistance to black Sigatoka.

| References | Enzymatic activity in <i>musa</i> spp. | Function |
|--------------------------------|---|---|
| Cruz-Martín et al., 2018 | APX - Ascorbate peroxidase | Antioxidant |
| | CHI - Chitinase GLU - β -1, 3-glucanase | Degradation of invading pathogen cell wall polysaccharides |
| | PAL - phenylalanine ammonia lyase | Synthesis of plant defense compounds, such as phytoalexins |
| | POX - Phenol peroxidase | Synthesis of lignin |
| | SOD - Superoxide dismutase | Oxidative stress due to increased production of H ₂ O ₂ |
| Torres et al., 2012 | CHI - Chitinase GLU - β -1, 3-glucanase | Degradation of invading pathogen cell wall polysaccharides |
| | PAL - phenylalanine ammonia lyase | Synthesis of plant defense compounds, such as phytoalexins |
| | H ₂ O ₂ - Peroxidase | Activates the plant's defense system |
| Mendoza-Rodríguez et al., 2017 | H ₂ O ₂ - Peroxidase | Activates the plant's defense system |
| Rodríguez et al., 2020 | H ₂ O ₂ - Peroxidase | Activates the plant's defense system |

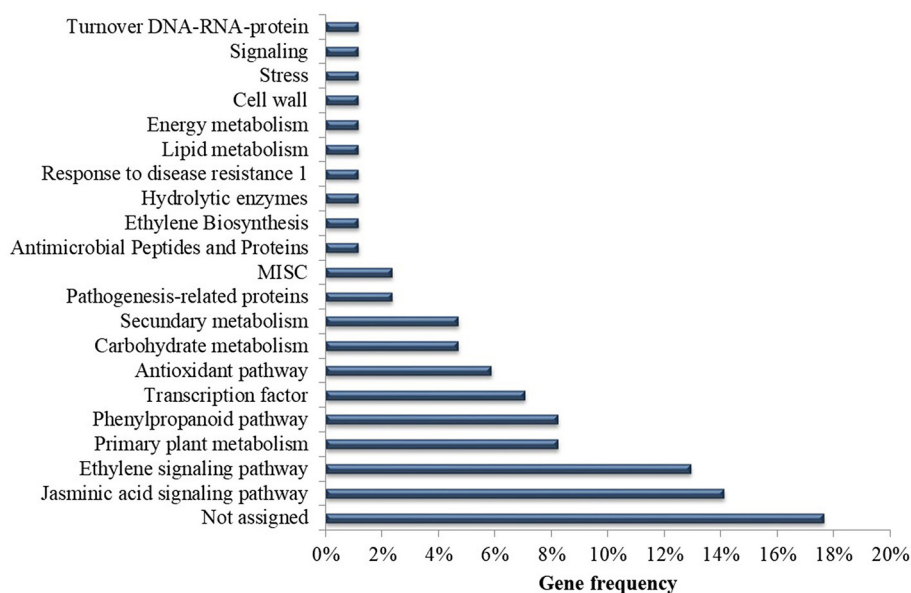


FIGURE 8 | Frequency of analyzed *Musa* genes according to predicted function in the selected publications recognized in the systematic review of genetic improvement of banana for resistance to black Sigatoka.

Swennen, 2004). IITA, together with NARO-Uganda, have also developed several East African cooking banana hybrids, known as NARITAs, which are high yielding and resistant to Black Sigatoka, with the most promising varieties now released to farmers (Ortiz, 2015). Hybrid plantains developed by IITA and considered resistant to *P. fijiensis*, known as PITAs, as well as resistant cooking hybrids, known as BITAs, are also available in countries such as Ghana, Ivory Coast, Cameroon, Uganda and Nigeria (Tenkouano and Swennen, 2004; Tenkouano et al., 2011). The EMBRAPA breeding program has also developed hybrid bananas that are currently being evaluated for resistance to *P. fijiensis* in Nigeria and Uganda by IITA, countries where bananas and plantains represent the principal food base (Amorim et al., 2021). Several hybrids developed by EMBRAPA also form the basis of banana production in Northern Brazil, a region widely affected by black Sigatoka.

Cultivars included in an ongoing adoption process by Brazilian producers include BRS Platina, a Prata-type, together with BRS Princesa, a Silk-type hybrid, which are both resistant to black Sigatoka, and *Fusarium oxysporum* f. sp. *cubense* race 1. In addition to the above, CIRAD and CARBAP also have advanced breeding programs that have also developed banana hybrids resistant to *P. fijiensis* through conventional strategies. These programs employ colchicine to duplicate chromosomes and aim to develop new cultivars rather than improving available germplasm (Tomepke et al., 2004). CARBAP maintains the largest collection of *Musa* spp. in the world, with more than 700 varieties from various geographic regions, and more than 150 banana cultivars (group AAB) susceptible to black Sigatoka (Mourichon et al., 1997; Tomepke et al., 2011). All of these breeding programs use crossing methods to obtain resistant materials.

Sources of Resistance

Black Sigatoka seriously affects dessert banana cultivars such as those of the Cavendish subgroup. One of the possible reasons for this susceptibility may lie in the monoculture format adopted for this subgroup, which theoretically may provide a favorable environment both for the emergence of resistance to fungicides within the pathogen population, as well as individuals with different virulence and/or aggressiveness characteristics (Churchill, 2011). A second reason is due to the type of host response to the pathogen. Although a reaction of the plant to attack by the pathogen has been recognized, the magnitude and development over time is regarded as insufficient to stop the progress of the fungus (Churchill, 2011; Torres et al., 2012; Cruz-Martín et al., 2018). For these reasons, one of the main recommended alternatives to fungicide-based approaches for the control of black Sigatoka is through the replacement of susceptible cultivars, such as those within the Cavendish subgroup, with agronomically appropriate resistant cultivars (Churchill, 2011).

In the present study, cultivars were reported with different levels of resistance, being classified as resistant, moderately resistant or susceptible to black Sigatoka. Among the most widely employed genotypes in the selected publications, the resistant *M. acuminata* wild diploid Calcutta 4, widely employed in breeding programs, and the susceptible triploid Grande Naine (Cavendish), used in commercial plantations for export and local consumption, both stand out in terms of frequency. Wild diploid *M. acuminata* bananas possess AA genomes and can harbor important sources of resistance genes for the genetic improvement of triploid cultivars (Timm et al., 2016). Studies into gene expression in Calcutta 4 have served as an approach to reveal candidate genes for resistance to the disease and to elucidate the mechanisms of resistance involved in the hypersensitivity response (Arango-Isaza et al., 2016; Timm et al., 2016; Mendoza-Rodríguez et al., 2017). In addition to Calcutta 4, other diploids resistant and moderately resistant to *P. fijiensis* include: Krasan Saichon, Zebrina, Birmanie, No. 118, Tuu Gia, PA Rayong, Pisang Cici, Malaccensis 1, 028003-01, Microcarpa, Pisang Lidi, Pisang Lilin, and Malbut. These have served as parentals for generation of improved diploids for subsequent introgression of genes in new cultivars (Nascimento et al., 2020).

Among the cultivars resistant to *P. fijiensis* mentioned in this review, BRS Maravilha, BRS Platina, FHIA-02, FHIA-18, and Galil 18 have adequate size and high yield potential, and represent alternatives to the traditional Prata subgroup. BRS Princesa, BRS Tropical and Caipira have also been promoted as alternatives to Silk bananas, with the cultivar Buccaneiro also an alternative to susceptible cultivars of the Gros Michel subgroup and appropriate for irrigated agrosystems (Weber et al., 2017).

The cultivar BRS Preciosa can also replace the commercial varieties Prata and Pacovan, without jeopardizing acceptability (Garruti et al., 2012; Amorim et al., 2021). In our review, we did not identify options of resistance to black Sigatoka in any cultivars of the Cavendish subgroup. Banana genetic improvement programs have, however, been focused on this

objective, with EMBRAPA, CIRAD and FHIA working on the development of pathogen resistant genotypes with similar fruit quality to Cavendish subgroup bananas.

Host Immune Responses to *P. fijiensis*

The identification of physical and chemical barriers related to banana defense has been the object of study to understand the mechanism of resistance to *P. fijiensis*. Lignification, together with production of phytoanticipins, phenols, phenylphenalenones, peroxidases, PAL (phenylalanine ammonia lyase), β -1,3 glucanase, and hydrogen peroxide all increase during incompatible interactions (Hoss et al., 2000; Otálvaro et al., 2007; Cruz-Cruz et al., 2010; Cavalcante et al., 2011; Torres et al., 2012; Sanchez-García et al., 2013; Hidalgo et al., 2016; Alakonya et al., 2018).

The sequencing of the reference genome of the diploid species *Musa acuminata* DH Pahang is an important resource for *Musa* improvement and has advanced understanding of banana evolution. In this study, numerous genes were identified that encode proteins potentially related to conserved components of PTI and ETI in monocots (D' Hont et al., 2012).

Analysis of gene expression is important for the identification of genes involved in plant-pathogen interactions. The genes C4H (cinnamate-4-hydroxylase), CHS (chalcone synthase), IRL (isoflavone reductase) and PAL (phenylalanine ammonia) are all related to the phenylpropanoid pathway. In the selected studies in this systematic review, these genes displayed similar up-regulated expression profiles in infected Calcutta 4 in contrast to an absence of such expression modulation in the susceptible cultivars Grande Naine (Mendoza-Rodríguez et al., 2018) and Williams (Alvarez et al., 2013), which presupposes recognition of the pathogen in Calcutta 4 and the appropriate expression of defense responses. The regulation of genes related to phytohormone defense responses is not entirely resolved in *Musa* spp. (Portal et al., 2011), although signaling associated with jasmonic acid (JA), salicylic acid (AS) and ethylene (ET) also participate in defense responses against pathogens. A total of 24 genes in the selected articles were related to signal transduction regulated by plant hormones, such as JA and ET (Supplementary Table 1). All genes related to the JA signaling pathway were found to be overexpressed in Calcutta 4 after inoculation with *P. fijiensis* (Rodríguez et al., 2020), whereas in the susceptible cultivar Williams, the activation of JA and ET defense responses was marginal, slow or non-existent, indicating potential suppression by pathogen effectors (Rodríguez et al., 2020). Pathogenesis-related proteins (PR) are induced in host plants after pathogen infection. PR-4 has been shown to have antifungal activity, disrupting cell polarity and binding to chitin in the cell wall of the fungus (Bormann et al., 1999; Portal et al., 2011). PR-10 exhibits ribonuclease and antifungal activity against pathogens in *Arachis hypogaea*, *Jatropha curcas*, and *Crocus sativus* (Chadha and Das, 2006; Gómez-Gómez et al., 2011; Agarwal et al., 2013). Here, in Calcutta 4, increased expression of genes encoding pathogenesis-related proteins PR-4 and PR-10 were found during interaction with *P. fijiensis* (Portal et al., 2011; Rodríguez et al., 2016). In a study by Mendoza-Rodríguez et al. (2018), gene expression in

the incompatible interaction in Calcutta 4 also reported positive regulation of the PSI gene (primary metabolism), TRX (an antioxidant) and SAMS (methyl cycle), suggesting roles in the defense response. In their work, negative regulation of genes from the phenylpropanoid pathway were also active in Grande Naine during initial phases of infection by *P. fijiensis*. Despite the advances in studies to date, further functional analyses of genes are warranted to validate use as candidate genes for resistance in susceptible banana cultivars (Timm et al., 2016). It is clear that there is no standardized protocol for studies of gene expression in banana during interaction with *P. fijiensis*, which may be a contributing factor to differences in results obtained.

Enzymes related to the defense response to *P. fijiensis* have been identified at different time points during infection and colonization. Raised enzymatic activities have been reported to occur earlier in certain resistant genotypes than in susceptible cultivars. As an example, Calcutta 4 showed a rapid induction of several defense-related enzymes, with peroxidase (POX), phenylalanine ammonia lyase (PAL), β -1, 3-glucanase (GLU) and well as the production of hydrogen peroxide (H_2O_2) during the first 72 h after inoculation, when compared to cv. Williams (Torres et al., 2012). H_2O_2 has been postulated to perform multiple functions in plant defense, with this reactive oxygen species involved in the rapid defense response of the plant identified as a hypersensitivity response (HR) (Awwad et al., 2019). One study has reported the accumulation of H_2O_2 associated with hypersensitivity reactions in Calcutta 4, enabling the rapid response in containing the development of the pathogen (Cavalcante et al., 2011). The enzymes POD and SOD are closely associated to oxidative stress responses caused by an increase in H_2O_2 . As such, increased activities in these enzymes, in addition to other antioxidant enzymes such as APX, have been described during incompatible responses (Cruz-Martín et al., 2018; Awwad et al., 2019). As the first enzyme in the phenylpropanoid pathway, the role of PAL in conversion of precursors in lignin biosynthesis has been well-elucidated. In relation to banana, however, its' role in the production of secondary metabolites such as phenylphenalenones and phytoanticipins, with potential activity against *P. fijiensis*, is poorly resolved (Hidalgo et al., 2009; Cruz-Cruz et al., 2010; Torres et al., 2012).

Study Environments

In relation to study environment, *in vitro* studies were conducted in a considerable proportion of the selected articles (45%). These comprised laboratory experiments investigating gene expression, enzymatic activity analysis, and gene function validation through transgenic approaches. Greenhouse studies corresponded to 24% of the articles, with focus on bioassays for evaluation of gene expression in *Musa* leaf tissues following inoculation with *P. fijiensis*. Field studies, which corresponded to only 16% of the articles, mostly focused on agronomic characterization and acceptance of resistant cultivars, with the exception of Barekye et al. (2011), who evaluated the contribution of diploid and tetraploid genotypes to triploid progenies, and Nascimento et al. (2020), who phenotyped 31 diploid accessions of Embrapa's germplasm collection for resistance.

Principal Techniques Employed

Evaluation of symptoms was described in 13% of the articles, with scales employed for measurement of black Sigatoka symptoms based on the quantification of percentage leaf area with characteristic lesions. In the selected articles, two different scales were cited: Fouré (1985), Alvarado-Capó et al. (2003) and Stover (1972), modified by Gauhl (1989). The main difference between the evaluation scales is that the former presents five evaluation stages for black Sigatoka in the greenhouse, whilst the latter describes six stages which can be used both in the greenhouse and in the field.

Amongst the techniques, one single study assessed surgical defoliation as a strategy for reducing disease severity (Jiménez and Brioso, 2018).

Transgenic approaches were also employed in 13% in the selected articles. Transformation protocols based on the use of fluorescent markers were employed with the pathogen to better understand the infection process in susceptible and resistant banana germplasm (Portal et al., 2012). Gene silencing strategies were also applied to determine gene function in the pathogen in relation to virulence (Onyilo et al., 2018). Vishnevetsky et al. (2011) developed a transformation system for improved tolerance to Sigatoka, with focus on endochitinase and styrene synthase candidate genes for resistance.

Hybridization and agronomic characterization represented only 9% of the frequency of the selected articles. The generation of banana triploids using this technique requires an understanding of the influence of the progenitors on potential resistance to black Sigatoka, as well as agronomic characteristics of the progenies generated (Barekye et al., 2011). Evaluation of growth and production of banana genotypes with resistance to *P. fijiensis* in comparison with cultivars susceptible to the disease was also carried out (Weber et al., 2017).

Among the biotechnological techniques employed, molecular markers such as gene-derived microsatellite markers have also been developed (Passos et al., 2012). These markers are appropriate for use in molecular genotyping and marker-assisted selection (MAS) in order to accelerate strategies for *Musa* genetic improvement.

LIMITATIONS OF THE REVIEW AND FUTURE RESEARCH

As this systematic review was highly specific to the *Musa* x *P. fijiensis* interaction with regard to genetic improvement for resistance, the number of studies was limited to only 24 articles suitable for inclusion. This indicates not only the need for further studies with this focus, but also that research trends may be focused more on other methods of controlling black Sigatoka in banana, such as those based on the use of fungicides or cultural control strategies for disease management.

Nevertheless, we strengthen as our closing remarks, that the banana genetic breeding for black Sigatoka based in the development of resistant cultivars through different methods is an efficient tool in the integrated management of the disease. It is possible, through genetic breeding to obtain basal, quantitative

resistance, since complete resistance has not yet been reported for the *Musa* x *P. fijiensis* pathosystem due to its complexity, especially as to selection of resistance genes with higher effect, and this is common for most agricultural crops (Kushalappa et al., 2016; Pilet-Nayel et al., 2017; Nascimento et al., 2020). Therefore, decreasing the symptoms of black Sigatoka obtained with limitations in the development of the pathogen in the tissues combined with cultural practices that aim reduction of the inoculum in the cultivated area is the best strategy for mitigating the impacts of the disease.

Banana possesses numerous characteristics that make genetic improvement a laborious and complex task. Despite this, breeding programs maintain a sustainable global banana agribusiness through the development of cultivars resistant to the main diseases of the crop. The process is inevitably slow, as *Musa* is a long cycle species that requires years for precise agronomic analysis of a new genotype to be completed. Agronomic studies combined with genetic studies employing biotechnological tools do, however, provide essential information for continuous genetic improvement.

The information contained in the literature on genes involved in the interaction between *Musa* x *P. fijiensis* is still relatively scarce, with the need for further focus on this pathosystem. Future advances in this direction will no doubt contribute to the elucidation of important processes occurring during this plant-pathogen interaction. In the short term, priorities for future studies are summarized below:

- In terms of accurate disease assessment, appropriate symptom scales are required that consider both greenhouse and field assessment, as symptomology can differ between these environments.
- Standardized inoculation protocols are recommended for the rapid screening of plants for resistance in greenhouse environments.
- Standardized protocols for analysis of gene expression in *Musa* during interaction with *P. fijiensis* are recommended, to reduce differences due to methodologies in results obtained by different research groups.
- The sources of resistance in *Musa* germplasm highlighted in the results are relevant for conventional breeding for development of disease resistant cultivars. No options for resistance to black Sigatoka were identified in any cultivars within the subgroup Cavendish.
- The development of a *Musa* x *P. fijiensis* interaction model at the molecular level is warranted, that infers how resistant genotype such as *M. acuminata* Calcutta 4 recognize the pathogen and develop a resistance response, as well as what types of weapons the pathogen launches to succeed in infection against susceptible genotypes.
- Gene editing based on CRISPR/Cas9 has been a recent major advance that can pave the way for large scale functional genomics, enabling validation and modification of candidate genes associated with characteristics such as resistance to biotic stresses (pathogens and pests) and tolerance to abiotic stresses (temperatures and extreme droughts). Although this approach has not yet been applied to the *Musa*-*P.*

fijiensis pathosystem, it offers considerable potential for the development of banana varieties with multiple and durable resistance and tolerance (Tripathi et al., 2019, 2020).

CONCLUSION

Invaluable tools and resources have been developed in recent years to further understand the interaction between *Musa* and *P. fijiensis*. These include reference genome sequences, bioinformatic tools, transcriptomic, proteomic, enzymatic, and histochemical data that have enabled identification of genes, proteins and intracellular events activated during pathogen invasion and host defense responses. Although breeding programs have developed hybrids resistant to *P. fijiensis*, the continued identification of additional sources of resistance is necessary, considering that resistance offered may have only a low durability, given the high variability of this fungus and potential appearance of aggressive pathogen variants.

The data collected in this systematic review highlight the considerable information accumulated in the last 10 years that is applicable for improvement of *Musa* for resistance to black Sigatoka. The *M. acuminata* genotype Calcutta 4 has been widely studied and can be a target for breeding programs and future studies. Certain questions can also be raised in relation to specific datasets highlighted here, such as which genes identified through expression studies as candidates for disease resistance are appropriate for transgenic or genetic editing systems, or which molecular markers are applicable in marker-assisted selection. The functional characterization of genes and proteins will advance understanding of function of these potential targets in the host, facilitating the development of novel disease control strategies.

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

JS, AR, FN, and EA developed the initial protocol for the development of the review. AS assisted in the use of the software. JS, AR, and FN carried out the data processing and development of the methodology, results, and discussion. EA, CF, FH, VA, and RM provided technical guidance and research supervision. All authors collaborated in the writing - revision, editing process, read, and agreed with the published version of the manuscript.

FUNDING

The authors would like to thank CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for the research productivity grants for EA and CF; Fapesb (Fundação de Amparo

à Pesquisa do Estado da Bahia) for providing Ph.D. scholarships to JS and AR; CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for providing Ph.D. scholarship to FN. The authors thank the Graduate Program in Biotechnology (PPGBiotec) of the State University of Feira de Santana.

REFERENCES

- Addy, H. S., Azizi, N. F., and Mihardjo, P. A. (2016). Detection of bacterial wilt pathogen and isolation of its bacteriophage from banana in lumajang area. *Int. J. Agron.* 2016, 1–7. doi: 10.1155/2016/5164846
- Agarwal, P., Bhatt, V., Singh, R., Das, M., Sopory, S. K., and Chikara, J. (2013). Pathogenesis-related gene, JcPR-10a, from *Jatropha curcas* exhibit RNase and antifungal activity. *Mol. Biotechnol.* 54, 412–425. doi: 10.1007/s12033-012-9579-7
- Alakonya, A. E., Kimunye, J., Mahuku, G., Amah, D., Uwimana, B., Brown, A., et al. (2018). Progress in understanding *Pseudocercospora* banana pathogens and the development of resistant *Musa* germplasm. *Plant Pathol.* 67, 759–770. doi: 10.1111/ppa.12824
- Alvarado-Capó, Y., Leiva-Mora, M., Dita-Rodríguez, M. A., Acosta, M., Cruz, M., Portal, N., et al. (2003). “Early evaluation of black leaf streak resistance by using mycelial suspensions of *Mycosphaerella fijiensis*,” in *Proceedings of the Workshop on Mycosphaerella Leaf Spot Diseases, San José*, eds L. Jacome, P. Lepoivre, D. Marin, R. Ortiz, R. Romero, and J. V. Escalant (Montpellier: International Network for the Improvement of Banana and Plantain), 169–175.
- Alvarez, J. C., Rodriguez, H. A., Rodriguez-Arango, E., Monsalve, Z. I., Morales, O. J. G., and Arango, I. R. E. (2013). Characterization of a differentially expressed phenylalanine ammonia-lyase gene from banana induced during *Mycosphaerella fijiensis* infection. *J. Plant Studies.* 2, 35–46. doi: 10.5539/jps.v2n2p35
- Amorim, E. P., Amorim, V. B. O., Silva, M. S., Haddad, F., Ferreira, C. F., and Santos-Serejo, J. A. (2021). “Developing hybrid banana varieties with improved properties,” in *Achieving Sustainable Cultivation of Bananas: Germplasm and Genetic Improvement*, Cambridge, eds G. H. J. Kema and A. Drenth (Burleigh Dodds Science Publishing), 1–17.
- Arango-Isaza, R. E., Diaz-Trujillo, C., Dhillon, B., Aerts, A., Carlier, J., Crane, C. F., et al. (2016). Combating a global threat to a clonal crop: banana black sigatoka pathogen *Pseudocercospora fijiensis* (Synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control. *PLoS Genet.* 12:e1005876. doi: 10.1371/journal.pgen.1005876
- Arinaitwe, I. K., Teo, C. H., Kayat, F., Tumuhimbise, R., Uwimana, B., Kubiriba, J., et al. (2019). Evaluation of banana germplasm and genetic analysis of an F1 population for resistance to *Fusarium oxysporum* f. sp. *cubense* race 1. *Euphytica* 215, 1–11. doi: 10.1007/s10681-019-2493-3
- Awwad, F., Bertrand, G., Grandbois, M., and Beaudoin, N. (2019). Reactive oxygen species alleviate cell death induced by thaxtomin A in *Arabidopsis thaliana* cell cultures. *Plants* 8:332. doi: 10.3390/plants8090332
- Bai, T. T., Xie, W. B., Zhou, P. P., Wu, Z. L., Xiao, W. C., Zhou, L., et al. (2013). Transcriptome and expression profile analysis of highly resistant and susceptible banana roots challenged with *Fusarium oxysporum* f. sp. *cubense* tropical race 4. *PLoS ONE* 8:e73945. doi: 10.1371/journal.pone.0073945
- Barekye, A., Tongoona, P., Derera, J., Laing, M. D., and Tushemereirwe, W. K. (2011). Contribution of synthetic tetraploids (AAAA) and diploids (AA) to black Sigatoka resistance and bunch weight to their triploid progenies. *Field Crops Res.* 122, 284–289. doi: 10.1016/j.fcr.2011.04.012
- Bebber, D. P. (2019). Climate change effects on Black Sigatoka disease of banana. *Philos Trans. R. Soc.* 374:2018026. doi: 10.1098/rstb.2018.0269
- Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* 60, 379–406. doi: 10.1146/annurev-arplant.57.032905.105346
- Bormann, C., Baier, D., Horr, I., Raps, C., Berger, J., Jung, G., et al. (1999). Characterization of a novel, antifungal, chitin-binding protein from *Streptomyces tendae* Tu901 that interferes with growth polarity. *J. Bacteriol.* 181, 7421–7429. doi: 10.1128/JB.181.24.7421-7429.1999
- Boutrot, F., and Zipfel, C. (2017). Function, discovery, and exploitation of plant pattern recognition receptors for broad-spectrum disease resistance. *Annu. Rev. Phytopathol.* 55, 257–286. doi: 10.1146/annurev-phyto-080614-120106
- Castañeda, N. E. N., Alves, G. S. C., Almeida, R. M., Amorim, E. P., Fortes Ferreira, C., Togawa, R. C., et al. (2017). Gene expression analysis in *Musa acuminata* during compatible interactions with *Meloidogyne incognita*. *Ann. Bot.* 119, 915–930. doi: 10.1093/aob/mcw272
- Cavalcante, M. J. B., Escoute, J., Madeira, J. P., Romero, R. E., Nicole, M., Oliveira, L. C., et al. (2011). Reactive oxygen species and cellular interactions between *Mycosphaerella fijiensis* and banana. *Trop. Plant Biol.* 4, 134–143. doi: 10.1007/s12042-011-9071-8
- Chadha, P., and Das, R. H. (2006). A pathogenesis related protein, AhPR10 from peanut: an insight of its mode of antifungal activity. *Planta* 225, 213–222. doi: 10.1007/s00425-006-0344-7
- Chisholm, S. T., Coaker, G., Day, B., and Staskawicz, B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124, 803–814. doi: 10.1016/j.cell.2006.02.008
- Chong, A. P. A. (2016). The Origin, Versatility and Distribution of Azole Fungicide Resistance in the Banana Black Sigatoka Pathogen *Pseudocercospora fijiensis*. thesis/PhD thesis, Wageningen, NL: Wageningen University. doi: 10.18174/387237
- Churchill, A. C. L. (2011). *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development, and the challenges of control. *Molec. Plant Pathol.* 12, 307–328. doi: 10.1111/j.1364-3703.2010.00672.x
- Coll, N. S., Eppe, P., and Dangl, J. L. (2011). Programmed cell death in the plant immune system. *Cell Death Differ.* 18, 1247–1256. doi: 10.1038/cdd.2011.37
- Creelman, R. A., and Mullet, J. E. (1995). Jasmonic acid distribution and action in plants: regulation during development and response to biotic and abiotic stress. *Proc. Natl. Acad. Sci. U.S.A.* 92, 4114–4119. doi: 10.1073/pnas.92.10.4114
- Cruz-Cruz, C. A., Ramírez-Tec, G., García-Sosa, K., Escalante-Erosa, F., Hill, L., Osbourn, A. E., et al. (2010). Phytoanticipins from banana (*Musa acuminata* cv. Grande Naine) plants, with antifungal activity against *Mycosphaerella fijiensis*, the causal agent of black Sigatoka. *Eur. J. Plant Pathol.* 126, 459–463. doi: 10.1007/s10658-009-9561-9
- Cruz-Martin, M., Acosta-Suárez, M., Mena, E., Roque, B., Pichardo, T., and Alvarado-Capó, Y. (2018). Effect of *Bacillus pumilus* CCIBP-C5 on *Musa-Pseudocercospora fijiensis* interaction. *3 Biotech.* 8:122. doi: 10.1007/s13205-018-1152-z
- Cui, H., Tsuda, K., and Parker, J. E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. *Annu. Rev. Plant Biol.* 66, 487–511. doi: 10.1146/annurev-arplant-050213-040012
- D’Hont, A., Denoeud, F., Aury, J. M., Baurens, F. C., Carreel, F., Garsmeur, O., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 21, 213–217. doi: 10.1038/nature11241
- Dangl, J. L., Horvath, D. M., and Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science* 341, 746–751. doi: 10.1126/science.1236011
- Dangl, J. L., and Jones, J. D. G. (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411, 826–833. doi: 10.1038/35081161
- de Costa Santos, C. M., de Mattos Pimenta, C. A., and Nobre, M. R., (2007). The PICO strategy for the research question construction and evidence search. *Rev. Lat. Am. Enfermagem.* 15, 508–11. doi: 10.1590/S0104-11692007000300023
- De Langhe, E., Vrydaghs, L., Maret, P., Perrier, X., and Denham, T. (2009). Why bananas matter: an introduction to the history of banana domestication. *Ethnobot. Res. Appl.* 7, 165–177. doi: 10.17348/era.7.0.165-177

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.657916/full#supplementary-material>

- Dita, M., Barquero, M., Heck, D., Mizubuti, E. S. G., and Staver, C. P. (2018). Fusarium wilt of banana: current knowledge on epidemiology and research needs toward sustainable disease management. *Front. Plant Sci.* 9:1468. doi: 10.3389/fpls.2018.01468
- Dong, X. (2001). Genetic dissection of systemic acquired resistance. *Curr. Opin. Plant Biol.* 4, 309–314. doi: 10.1016/S1369-5266(00)00178-3
- Emediato, F. L., Nunes, F. A. C., de Camargo Teixeira, C., Passos, M. A. N., Bertoli, D. J., Pappas Jr, G. J., et al. (2009). "Characterization of resistance gene analogs in *Musa acuminata* cultivars contrasting in resistance to biotic stresses," in *Induced Plant Mutations in the Genomics Era*, ed Q. Y. Shu (Rome: Food and Agriculture Organization of the United Nations), 443–445.
- Emediato, F. L., Passos, M. A. N., de Camargo Teixeira, C., Pappas Jr, G. J., and Miller, R. N. G. (2013). Analysis of expression of NBS-LRR resistance gene analogs in *Musa acuminata* during compatible and incompatible interactions with *Mycosphaerella musicola*. *Acta Hort.* 986, 255–258. doi: 10.17660/ActaHortic.2013.986.28
- Falcomer, A. L., Riquette, R. F. R., de Lima, B. R., Ginani, V. C., and Zandonadi, R. P. (2019). Health benefits of green banana consumption: a systematic review. *Nutrients* 11:1222. doi: 10.3390/nu11061222
- FAOSTAT (2021). *FAOSTAT Online Database*. Available at <http://faostat.fao.org/> (accessed January 2021).
- Felix, G., Duran, J. D., Volko, S., and Bolter, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J.* 18, 265–276. doi: 10.1046/j.1365-313x.1999.00265.x
- Ferreira, C. F., Silva, S. O., Sobrinho, N. P. D., and Paz, O. P. (2004). Molecular characterization of banana (AA) diploids with contrasting levels of black and yellow sigatoka resistance. *Am. J. Appl. Sci.* 1, 276–278. doi: 10.3844/ajassp.2004.276.278
- Fouré, E. (1985). *Black Leaf Streak Disease of Bananas and Plantains (Mycosphaerella fijiensis Morelet). Study of the Symptoms and Stages of the Disease in Gabon*. Paris: CIRAD-IRFA.
- Friesen, T. L. (2016). Combating the Sigatoka disease complex on banana. *PLoS Genet.* 12:e1006234. doi: 10.1371/journal.pgen.1006234
- Garruti, D. D. S., Matias, M. D. L., Facundo, H. V. D. V., Silva, E. D. O., Costa, J. N. D., and Silva, M. A. A. P. D. (2012). Acceptance of banana cultivars resistant to Black Sigatoka by the consumer market of Northeast Brazilian region. *Cienc. Rural.* 42, 948–954. doi: 10.1590/S0103-84782012000500030
- Gauhl, F. (1989). Untersuchungen zur Epidemiologie und Ökologie der Schwarzen Sigatoka-Krankheit (*Mycosphaerella fijiensis* Morelet) an Kochbananen (*Musa* sp.) in Costa Rica. thesis/PhD. Thesis. Göttingen: University of Göttingen
- Gómez-Gómez, L., Rubio-Moraga, A., and Ahrazem, O. (2011). Molecular cloning and characterization of a pathogenesis-related protein CsPR10 from *Crocus sativus*. *Plant Biol.* 13, 297–303. doi: 10.1111/j.1438-8677.2010.00359.x
- Gong, B. Q., Wang, F. Z., and Li, J. F. (2020). Hide-and-seek: chitin-triggered plant immunity and fungal counterstrategies. *Trends Plant Sci.* 25, 805–816. doi: 10.1016/j.tplants.2020.03.006
- Gururani, M. A., Venkatesh, J., Upadhyaya, C. P., Nookaraju, A., Pandey, S. K., and Park, S. W. (2012). Plant disease resistance genes: Current status and future directions. *Physiol. Mol. Plant Pathol.* 78, 51–65. doi: 10.1016/j.pmpp.2012.01.002
- Hidalgo, W., Chandran, J. N., Menezes, R. C., Otálvaro, F., and Schneider, B. (2016). Phenylphenalenones protect banana plants from infection by *Mycosphaerella fijiensis* and are deactivated by metabolic conversion. *Plant Cell Environ.* 39, 492–513. doi: 10.1111/pce.12630
- Hidalgo, W., Duque, L., Saez, J., Arango, R., Gil, J., Rojano, B., et al. (2009). Structure-activity relationship in the interaction of substituted perinaphthenones with *Mycosphaerella fijiensis*. *J. Agric. Food Chem.* 57, 7417–7421. doi: 10.1021/jf901052e
- Hoss, R., Helbig, J., and Bochow, H. (2000). Function of host and fungal metabolites in resistance response of banana and Plantain in the black Sigatoka disease pathosystem (*Musa* spp.) (*Mycosphaerella fijiensis*). *J. Phytopathol.* 148, 387–394. doi: 10.1046/j.1439-0434.2000.00530.x
- Jiménez, J. L. S., and Brioso, P. S. T. (2018). Surgery or surgical defoliation in Grand Naine banana in the control of black Sigatoka in the state of Rio de Janeiro. *Rev. Bras. Frutic.* 40:5. doi: 10.1590/0100-29452018144
- Jimenez, M., Van der Veken, L., Neirynck, H., Rodríguez, H., Ruiz, O., and Swennen, R. (2007). Organic banana production in Ecuador: Its implications on black Sigatoka development and plant–soil nutritional status. *Renew. Agric. Food Syst.* 22, 297–306. doi: 10.1017/S1742170507001895
- Jones, A. T., O'Connell, N. K., and David, A. S. (2020). Epidemiology of functional stroke mimic patients: a systematic review and meta-analysis. *Eur. J. Neurol.* 27, 18–26. doi: 10.1111/ene.14069
- Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Kushalappa, A. C., Yogendra, K. N., and Karre, S. (2016). Plant innate immune response: qualitative and quantitative resistance. *Crit. Rev. Plant Sci.* 35, 38–55. doi: 10.1080/07352689.2016.1148980
- Leiva-Mora, M., Alvarado-Capó, Y., Acosta-Suárez, M., Cruz-Martín, M., Sánchez-García, C., and Roque, B. (2010). Protocolo para la inoculación artificial de plantas de *Musa* spp. con *Mycosphaerella fijiensis* y evaluación de su respuesta mediante variables epifitológicas y componentes de la resistencia. *Biot. Veg.* 10, 79–88.
- Li, B., Meng, X., Shan, L., and He, P. (2016). Transcriptional regulation of pattern-triggered immunity in plants. *Cell Host Microbe.* 19, 641–650. doi: 10.1016/j.chom.2016.04.011
- Li, C. Y., Deng, G. M., Yang, J., Viljoen, A., Jin, Y., Kuang, R. B., et al. (2012). Transcriptome profiling of resistant and susceptible Cavendish banana roots following inoculation with *Fusarium oxysporum* f. sp. *cubense* tropical race 4. *BMC Genom.* 13:374. doi: 10.1186/1471-2164-13-374
- Mendoza-Rodríguez, M., Ocaña, B., Acosta-Suárez, M., Roque, B., Hernández, M., and Cruz-Martín, M. (2017). Effect of H₂O₂ application during 'Grande naine'–*Mycosphaerella fijiensis* interaction. *Biot. Veg.* 17, 153–159.
- Mendoza-Rodríguez, M. F. (2014). Avances en los estudios sobre la interacción *Musa* spp.–*Mycosphaerella fijiensis* morelet. *Biot. Veg.* 14, 131–137.
- Mendoza-Rodríguez, M. F., Portal, O., Oloriz, M. I., Ocaña, B., Rojas, L. E., Acosta-Suárez, M., et al. (2018). Early regulation of primary metabolism, antioxidant, methyl cycle and phenylpropanoid pathways during the *Mycosphaerella fijiensis*–*Musa* spp. interaction. *Trop. Plant Pathol.* 43, 1–9. doi: 10.1007/s40858-017-0188-7
- Miller, R. N., Bertoli, D. J., Baurens, F. C., Santos, C. M., Alves, P. C., Martins, N. F., et al. (2008). Analysis of non-TIR NBS-LRR resistance gene analogs in *Musa acuminata* Colla: isolation, RFLP marker development, and physical mapping. *BMC Plant Biol.* 8:15. doi: 10.1186/1471-2229-8-15
- Miller, R. N. G., Passos, M. A. N., Emediato, F. L., Teixeira, C. C., and Pappas Júnior, G. J. (2011). Candidate resistance gene discovery: resistance gene analog characterization and differential gene expression analysis in *Musa*–*Mycosphaerella* host–pathogen interactions. *Acta Hort.* 897, 179–185. doi: 10.17660/ActaHortic.2011.897.19
- Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Curr. Opin. Plant Biol.* 15, 349–357. doi: 10.1016/j.pbi.2012.05.006
- Mourichon, X., Carlier, J., and Fouré E. (1997). *Les Cercosporioses: Maladies des Raies Noires (Cercosporiose Noire), Maladie de Sigatoka (Cercosporiose Jaune) in Maladie des Musa. Fiche Technique n°8*. Montpellier: INIBAP, Parc Scientifique Agropolis.
- Nascimento, F. S., Sousa, Y. M., Rocha, A. J., Ferreira, C. F., Haddad, F., and Amorim, E. P. (2020). Sources of black Sigatoka resistance in wild banana diploids. *Rev. Bras. Frutic.* 42:4. doi: 10.1590/0100-29452020038
- Nowakunda, K., Barekye, A., Ssali, R. T., Namaganda, J., Tushemereirwe, W. K., Nabulya, G., et al. (2015). Kiwangaazi (syn KABANA 6H) Black Sigatoka nematode and banana weevil tolerant Matooke hybrid banana released in Uganda. *HortScience* 50, 621–623. doi: 10.21273/HORTSCI.50.4.621
- Oiram-Filho, F., Lopes, M. M. A., Matias, M. L., Braga, T. R., Aragão, F. A. S., Silveira, M. R. S., et al. (2019). Shelf-life estimation and quality of resistant bananas to black leaf streak disease during ripening. *Sci. Hortic.* 251, 267–275. doi: 10.1016/j.scienta.2019.03.029
- Onyilo, F., Tusiime, G., Tripathi, J. N., Chen, L. H., Falk, B., Stergiopoulos, L., et al. (2018). Silencing of the mitogen-activated protein kinases (MAPK) Fus3 and Slt2 in *Pseudocercospora fijiensis* reduces growth and virulence on host plants. *Front. Plant Sci.* 9:291. doi: 10.3389/fpls.2018.00291
- Ortiz, R. R. (2015). *Plant Breeding in the Omics Era*. Springer International Publishing.
- Otálvaro, F., Nanclares, J., Vázquez, L. E., Quinones, W., Echeverri, F., Arango, R., et al. (2007). Phenalenone-type compounds from *Musa acuminata*

- var. “Yangambi km 5” (AAA) and their activity against *Mycosphaerella fijiensis*. *J. Nat. Prod.* 70, 887–890. doi: 10.1021/np070091e
- Passos, M. A. N., Cruz, V. O., Emediato, F. L., Teixeira, C. C., Azevedo, V. C. R., Brasileiro, A. C. M., et al. (2013). Analysis of the leaf transcriptome of *Musa acuminata* during interaction with *Mycosphaerella musicola*: gene assembly, annotation and marker development. *BMC Genom.* 14:78. doi: 10.1186/1471-2164-14-78
- Passos, M. A. N., Cruz, V. O., Emediato, F. L., Teixeira, C. C., Souza Jr, M. T., Matsumoto, T., et al. (2012). Development of expressed sequence tag and expressed sequence tag-simple sequence repeat marker resources for *Musa acuminata*. *AoB Plants* 2012:pls030. doi: 10.1093/aobpla/pls030
- Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., et al. (2017). Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Front. Plant Sci.* 8:1838. doi: 10.3389/fpls.2017.01838
- Portal, O., Acosta-Suárez, M., Ocaña, B., Schäfer, W., Jiménez, E., and Höfte, M. (2012). A green fluorescent protein-transformed *Mycosphaerella fijiensis* strain shows increased aggressiveness on banana. *Austr. Plant Pathol.* 41, 645–647. doi: 10.1007/s13313-012-0155-1
- Portal, O., Izquierdo, Y., Vleeschauwer, D., Sánchez-Rodríguez, A., Mendoza-Rodríguez, M., Acosta-Suárez, M., et al. (2011). Analysis of expressed sequence tags derived from a compatible *Mycosphaerella fijiensis*–banana interaction. *Plant Cell Rep.* 30, 913–928. doi: 10.1007/s00299-011-1008-z
- Rodríguez, H. A., Hidalgo, W. F., Sanchez, J. D., Menezes, R. C., Schneider, B., Arango, R. E., et al. (2020). Differential regulation of jasmonic acid pathways in resistant (Calcutta 4) and susceptible (Williams) banana genotypes during the interaction with *Pseudocercospora fijiensis*. *Plant Pathol.* 69, 872–882. doi: 10.1111/ppa.13165
- Rodríguez, H. A., Rodríguez-Arango, E., Morales, J. G., Kema, G., and Arango, R. E. (2016). Defense gene expression associated with biotrophic phase of *Mycosphaerella fijiensis* M. Morelet infection in banana. *Plant Dis.* 100, 1170–1175. doi: 10.1094/PDIS-08-15-0950-RE
- Rodríguez-García, C. M., Canché-Gómez, A. D., Sáenz-Carbonell, L., Peraza-Echeverría, L., Canto-Canché, B., Islas-Flores, I., et al. (2016). Expression of MfAvr4 in banana leaf sections with black leaf streak disease caused by *Mycosphaerella fijiensis*: a technical validation. *Austr. Plant Pathol.* 45, 481–488. doi: 10.1007/s13313-016-0431-6
- Sanchez-García, Y. A. C., Acosta-Suárez, M., Leiva-Mora, M., Cruz-Martín, M., and Roque, B. (2013). Quantification of phenols in lesions caused by *Mycosphaerella fijiensis* Morelet in ‘Cavendish Naine’. *Rev. Protección Veg.* 28, 149–152.
- Santos, A. S., Amorim, E. P., Ferreira, C. F., and Pirovani, C. P. (2018). Water stress in *Musa* spp.: a systematic review. *PLoS ONE* 13:e0208052. doi: 10.1371/journal.pone.0208052
- Seenivasan, N. (2017). Identification of burrowing nematode (*Radopholus similis*) resistance in banana (*Musa* spp.) genotypes for natural and challenge inoculated populations. *Arch. Phytopathol. Pflanzenschutz* 50, 415–437. doi: 10.1080/03235408.2017.1321871
- Simmonds, N. W., and Shepherd, K. (1955). The taxonomy and origins of the cultivated bananas. *J. Linn. Soc. Bot.* 55, 302–312. doi: 10.1111/j.1095-8339.1955.tb00015.x
- Spoel, S. H., and Dong, X. (2012). How do plants achieve immunity? Defence without specialized immune cells. *Nat. Rev. Immunol.* 12, 89–100. doi: 10.1038/nri3141
- Stover, R. H. (1972). *Banana Plantain and Abaca Diseases*. Toronto, ON: Commonwealth Mycological Institute.
- Sullivan, K. S. U. S., Pillay, M., and Yasmina, J. F. (2012). Polymorphism at selected defence gene analogs (DGA) of *Musa* accessions in Mauritius. *Afr. J. Biotechnol.* 11, 11207–11220. doi: 10.5897/AJB11.2211
- Sun, D., Hu, Y., Zhang, L., Mo, Y., and Xie, J. (2009). Cloning and analysis of Fusarium wilt resistance gene analogs in ‘Goldfinger’ banana. *Mol. Plant Breed.* 7, 1215–1222. doi: 10.5376/mp.2010.01.0001
- Tenkouano, A., Pillay, M., and Coulbaly, O. (2011). “Hybrid distribution to farmers and adoption and challenges,” in *Banana Breeding: Constraints and Progress*, eds M. Pillay and A. Tenkouano (Boca Raton, FL: CRC Press), 305–319.
- Tenkouano, A., and Swennen, R. (2004). Progress in breeding and delivering improved plantain and banana to African farmers. *Chron. Hort.* 44, 9–15.
- Thomma, B. P. H. J., Nürnberger, T., and Joosten, M. H. A. J. (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *Plant Cell.* 23, 4–15. doi: 10.1105/tpc.110.082602
- Timm, E. S., Pardo, L. H., Coello, R. P., Navarrete, T. C., Villegas, O. N., and Ordóñez, E. S. (2016). Identification of differentially-expressed genes in response to *Mycosphaerella fijiensis* in the resistant *Musa* accession ‘Calcutta-4’ using suppression subtractive hybridization. *PLoS ONE* 11:e0160083. doi: 10.1371/journal.pone.0160083
- Tomekpe, K., Kwa, M., Dzomeku, B. M., and Ganry, J. (2011). CARBAP and innovation on the plantain banana in Western and Central Africa. *Int. J. Agric. Sustain.* 9, 264–273. doi: 10.3763/ijas.2010.0565
- Tomekpe, K., Jenny, C., and Escalant, J. V. (2004). A review of conventional improvement strategies for *Musa*. *InfoMusa* 13, 2–6.
- Torres, J. M., Calderón, H., Rodríguez-Arango, E., Morales, J. G., and Arango, R. (2012). Differential induction of pathogenesis-related proteins in banana in response to *Mycosphaerella fijiensis* infection. *Eur. J. Plant Pathol.* 133, 887–898. doi: 10.1007/s10658-012-0012-7
- Tripathi, L., Ntui, V. O., and Tripathi, J. N. (2019). Application of genetic modification and genome editing for developing climate-smart banana. *Food Energy Secur.* 8:e00168. doi: 10.1002/fes3.168
- Tripathi, L., Ntui, V. O., and Tripathi, J. N. (2020). CRISPR/Cas9-based genome editing of banana for disease resistance. *Curr. Opin. Plant Biol.* 56, 118–126. doi: 10.1016/j.pbi.2020.05.003
- Tumuhimbise, R., Barekye, A., Kubiriba, J., Akankwasa, K., Arinaitwe, I. K., Karamura, D., et al. (2018). New high-yield cooking banana cultivars with multiple resistances to pests and diseases (‘NAROBan1’, ‘NAROBan2’, ‘NAROBan3’, and ‘NAROBan4’) released in Uganda. *HortScience* 53, 1387–1389. doi: 10.21273/HORTSCI13207-18
- Valmayor, R. V. (2001). Classification and characterization of *Musa exotica*, *M. alisanaya* and *M. acuminata* ssp. errans. *Infomusa* 10, 35–39.
- Vishnevetsky, J., White, T. L., Palmateer, A. J., Flaishman, M., Cohen, Y., Elad, Y., et al. (2011). Improved tolerance toward fungal diseases in transgenic Cavendish banana (*Musa* spp. AAA group) cv. Grand Nain. *Transgenic Res.* 20, 61–72. doi: 10.1007/s11248-010-9392-7
- Wan, J., Zhang, S., and Stacey, G. (2004). Activation of a mitogen-activated protein kinase pathway in Arabidopsis by chitin. *Mol. Plant Pathol.* 5, 125–135. doi: 10.1111/j.1364-3703.2004.00215.x
- Wang, Z., Zhang, J., Jia, C., Liu, J., Li, Y., Yin, X., et al. (2012). De Novo characterization of the banana root transcriptome and analysis of gene expression under *Fusarium oxysporum* f. sp. *cubense* tropical race 4 infection. *BMC Genom.* 13:650. doi: 10.1186/1471-2164-13-650
- Weber, O. B., Garruti, D. D. S., and Norões, N. P. (2017). Performance of banana genotypes with resistance to black leaf streak disease in Northeastern Brazil. *Pesqui. Agropecu. Bras.* 52, 161–169. doi: 10.1590/s0100-204x2017000300003
- Wright, R. W., Brand, R. A., Dunn, W., and Spindler, K. P. (2007). How to write a systematic review. *Clin. Orthop. Relat. Res.* 455, 23–29. doi: 10.1097/BLO.0b013e31802c9098
- Yonow, T., Ramirez-Villegas, J., Abadie, C., Darnell, R. E., Ota, N., and Kriticos, D. J. (2019). Black Sigatoka in bananas: Ecoclimatic suitability and disease pressure assessments. *PLoS ONE* 14:e0220601. doi: 10.1371/journal.pone.0220601
- 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 Soares, Rocha, Nascimento, Santos, Miller, Ferreira, Haddad, Amorim and Amorim. 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.



Introgression of Maize Lethal Necrosis Resistance Quantitative Trait Loci Into Susceptible Maize Populations and Validation of the Resistance Under Field Conditions in Naivasha, Kenya

OPEN ACCESS

Edited by:

Harsh Raman,
New South Wales Department of
Primary Industries, Australia

Reviewed by:

Bernardo Ordas,
Consejo Superior de Investigaciones
Científicas (CSIC), Spain
Francis Chuks Ogbonnaya,
Grains Research and Development
Corporation, Australia

*Correspondence:

Pangirayi B. Tongoona
ptongoona@wacci.ug.edu.gh;
tongoona@ukzn.ac.za
MacDonald Bright Jumbo
b.jumbo@cgjar.org

Specialty section:

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

Received: 04 January 2021

Accepted: 29 March 2021

Published: 03 May 2021

Citation:

Awata LAO, Ifie BE, Danquah E,
Jumbo MB, Suresh LM, Gowda M,
Marchelo-Dragna PW, Olsen MS,
Shorinola O, Yao NK,
Boddupalli PM and
Tongoona PB (2021) Introgression of
Maize Lethal Necrosis Resistance
Quantitative Trait Loci Into
Susceptible Maize Populations and
Validation of the Resistance Under
Field Conditions in Naivasha, Kenya.
Front. Plant Sci. 12:649308.
doi: 10.3389/fpls.2021.649308

Luka A. O. Awata¹, Beatrice E. Ifie², Eric Danquah², MacDonald Bright Jumbo^{3*},
L. Mahabaleswara Suresh⁴, Manje Gowda⁴, Philip W. Marchelo-Dragna⁵,
Michael Scott Olsen⁴, Oluwaseyi Shorinola^{6,7}, Nasser Kouadio Yao⁶, Prasanna
M. Boddupalli⁴ and Pangirayi B. Tongoona^{2*}

¹Directorate of Research, Ministry of Agriculture and Food Security, Juba, South Sudan, ²West Africa Centre for Crop Improvement (WACCI), College of Basic and Applied Sciences, University of Ghana, Legon, Ghana, ³International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Bulawayo, Zimbabwe, ⁴International Maize and Wheat Improvement Center (CIMMYT), World Agroforestry Centre (ICRAF), Nairobi, Kenya, ⁵Department of Agricultural Sciences, College of Natural Resources and Environmental Studies, University of Juba, Juba, South Sudan, ⁶Biosciences eastern and central Africa (BeCA) Hub, International Livestock Research Institute (ILRI), Nairobi, Kenya, ⁷John Innes Centre, Norwich, United Kingdom

Maize lethal necrosis (MLN), resulting from co-infection by maize chlorotic mottle virus (MCMV) and sugarcane mosaic virus (SCMV) can cause up to 100% yield losses in maize in Africa under serious disease conditions. Maize improvement through conventional backcross (BC) takes many generations but can significantly be shortened when molecular tools are utilized in the breeding process. We used a donor parent (KS23-6) to transfer quantitative trait loci (QTL) for resistance to MLN into nine adapted but MLN susceptible lines. Nurseries were established in Kiboko, Kenya during 2015–2017 seasons and BC₃F₂ progeny were developed using marker assisted backcrossing (MABC) approach. Six single nucleotide polymorphism (SNP) markers linked to QTL for resistance to MLN were used to genotype 2,400 BC₃F₂ lines using Kompetitive Allele Specific PCR (KASP) platform. We detected that two of the six QTL had major effects for resistance to MLN under artificial inoculation field conditions in 56 candidate BC₃F₂ lines. To confirm whether these two QTL are reproducible under different field conditions, the 56 BC₃F₂ lines including their parents were evaluated in replicated trials for two seasons under artificial MLN inoculations in Naivasha, Kenya in 2018. Strong association of genotype with phenotype was detected. Consequently, 19 superior BC₃F₂ lines with favorable alleles and showing improved levels of resistance to MLN under artificial field inoculation were identified. These elite lines represent superior genetic resources for improvement of maize hybrids for resistance to MLN. However, 20 BC₃F₂ lines were fixed for both KASP markers but were susceptible to MLN under field conditions, which could suggest weak linkage between the KASP

markers and target genes. The validated two major QTL can be utilized to speed up the breeding process but additional loci need to be identified between the KASP markers and the resistance genes to strengthen the linkage.

Keywords: maize, backcross, kompetitive allele specific PCR, alleles, maize lethal necrosis, introgression

INTRODUCTION

Maize is the major food crop in Sub-Saharan Africa, however, its productivity remains low due to various production constraints such as biotic and abiotic factors. Recently, maize lethal necrosis (MLN) has emerged as one of the most deadly maize diseases in the region with high yield losses. MLN is caused by co-infection of maize plants by Maize chlorotic mottle virus (MCMV) in combination with any of the cereal viruses in the family *Potyviridae*, such as Sugarcane mosaic virus (SCMV), Maize dwarf mosaic virus (MDMV), or Wheat streak mosaic virus (WSMV). In east Africa, it has been established that it is mostly SCMV in combination with MCMV causing MLN. MLN disease causing viruses are transmitted by vectors such as thrips and beetles for MCMV (Nault et al., 1978), and aphids for potyviruses such as SCMV and MDMV (Brandes, 1920; Pemberton and Charpentier, 1969). MLN can cause losses in maize ranging from 30 to 100% depending on disease pressure (Gowda et al., 2018; Awata et al., 2019). Elite maize lines used in countries like South Sudan and others are highly susceptible to MLN and farmers risk losing their crops and money if MLN is not controlled. Therefore, efforts to develop high yielding varieties with resistance to MLN are urgently required. Breeding for host resistance can provide added advantage to farmers in terms of costs as compared to spraying against the vectors using chemicals, which is expensive and results in pollution to the environment. Conventional backcrossing is a routine breeding approach used for introgression of novel genes into the genetic backgrounds of adapted but susceptible germplasm but requires 8–10 generations to develop lines with desired characteristics. Studies to identify genomic regions associated with MLN resistance using linkage mapping revealed three major quantitative trait loci (QTL) on chromosomes 3, 6, and 9 that were consistently detected in at least two populations (Gowda et al., 2018) with recessive genetic effects. Introgression of genes for MLN resistance into the adapted lines using molecular markers is a quick option for fast-tracking development of varieties with resistance to MLN. Marker assisted backcrossing (MABC) has been widely used in improvement of maize for traits of economic importance including resistance to diseases (Lübberstedt et al., 2006; Muthusamy et al., 2014; Feng et al., 2015; Rasheed et al., 2016). Therefore, use of MABC can be used to speed up identification of fixed QTL conferring resistance to MLN into the backgrounds of adapted but susceptible maize lines. Kompetitive allele specific PCR (KASP) markers, developed by LGC Genomics (Teddington, United Kingdom),¹ is a PCR-based homogeneous fluorescent single nucleotide polymorphism (SNP) genotyping system. It has the power to detect single nucleotide polymorphism at a

specific locus using dual Fluorescent Resonance Energy Transfer (FRET; Semagn et al., 2014). KASP has high throughput, low cost, and more robust than other genotyping assays such as Restriction Fragment Length Polymorphism (RFLP), Randomly Amplified Polymorphic DNA (RAPD), and Amplified Fragment Length Polymorphism (AFLP), which require longer time and have higher cost per sample. KASP technology has been utilized on various crops including wheat (Rasheed et al., 2016) and cordgrass (Graves et al., 2016). The objectives of this study were: (i) introgression of MLN resistance from a resistant source into adapted but susceptible elite maize lines using the KASP method; (ii) validate the effect of the introgressed QTL for resistance to MLN in lines evaluated under MLN artificial inoculation in the field; and (iii) identify resistant lines for future breeding.

MATERIALS AND METHODS

Genetic Materials

Genetic materials were provided by CIMMYT Global Maize Program and consisted of two maize categories: (i) MLN resistant donor line (KS23-6) developed by Kasetsart University in Thailand, which is a yellow maize line and considered suitable parent for maize improvement in Africa because of its tropical adaptation; (ii) 19 elite but MLN susceptible white CIMMYT lines, with diverse tropical backgrounds and each line belonging to one of the two heterotic groups (A and B) and are commonly used for hybrid development in the region including South Sudan due to their high yield performance and resistance to major foliar diseases (Supplementary Appendix 1).

Development of Bi-Parental Backcross Populations

Crossing blocks to develop backcross (BC) populations were established at CIMMYT in Kiboko, between July 2015 and 2017 cropping seasons. Kiboko is located within the dry-mid altitude environment at 37° 75'E and 20 15'S, and 975 masl in Makueni County, Kenya, with mean temperature ranging from 14.3 to 35.1°C (Ziyomo and Bernardo, 2013; Odiyo et al., 2014). The first crossing block was set up in April 2015 and bi-parental populations were formed by crossing KS23-6 as pollen donor to the 19 selected elite but MLN susceptible lines (Awata et al., 2018). Adequate moisture was supplied through drip irrigation and standard agronomic practices and nursery management were applied. A nursery to develop backcross populations was established in October 2015, where 19 F₁ populations were grown in single-row plots of 4.0 m spaced at 0.75 × 0.25 m (Table 1).

¹www.lgcgenomics.com

TABLE 1 | List of 19 bi-parental crosses generated and used to develop BC₃F₂ lines genotyped for resistance to MLN using two polymorphic SNP markers linked to major QTL for resistance to MLN in BecA-ILRI Hub Lab in July 2017.

| SN | Population | Bi-parental cross | SN | Population | Biparental cross |
|----|------------|--------------------|----|------------|--------------------|
| 1 | Pop1 | KS23-6 × CML567 | 11 | Pop11 | KS23-6 × CML547 |
| 2 | Pop2 | KS23-6 × CML568 | 12 | Pop12 | KS23-6 × CML566 |
| 3 | Pop3 | KS23-6 × CML442 | 13 | Pop13 | KS23-6 × CML569 |
| 4 | Pop4 | KS23-6 × CML537 | 14 | Pop14 | KS23-6 × CML570 |
| 5 | Pop5 | KS23-6 × CML548 | 15 | Pop15 | KS23-6 × CKL05017 |
| 6 | Pop6 | KS23-6 × CML572 | 16 | Pop16 | KS23-6 × CKL05019 |
| 7 | Pop7 | KS23-6 × CKDHL0106 | 17 | Pop17 | KS23-6 × CML539 |
| 8 | Pop8 | KS23-6 × CKDHL0323 | 18 | Pop18 | KS23-6 × CML540 |
| 9 | Pop9 | KS23-6 × CML444 | 19 | Pop19 | KS23-6 × CKDHL0186 |
| 10 | Pop10 | KS23-6 × CML511 | 20 | KS23-6* | |

*Donor parent for resistance to MLN.

Larger population size was required so that both major QTL associated with resistance to MLN could be detected (Bouchez et al., 2002; Hospital, 2003; Vales et al., 2005; Ribaut and Ragot, 2007). Therefore, 10 F₁ individual plants (females) were tagged per population and each backcrossed to its recurrent parent (males), hence BC₁F₁ populations were developed (Brown and Caligari, 2008; Acquaaah, 2012). The BC₁F₁ progeny were planted in March 2016. About 60 agronomically healthy BC₁F₁ individuals (females) were labeled within each population and backcrossed to their respective recurrent parents (males), to generate BC₂F₁. The BC₂F₁ populations were planted in the nursery in August 2016 and about 60 clean plants in each population were tagged and backcrossed to the recurrent parents where BC₃F₁ populations were generated. The BC₃F₁ populations were planted in January 2017 and tissue samples were collected from this nursery and sent to LGC in United Kingdom and genotyped with 100 markers for MLN and 250–300 individuals per population were selected based on the marker information received from LGC. The selected 250–300 individuals for each population were manually self-pollinated and BC₃F₂ seeds obtained (Table 2). The BC₃F₂ populations were eventually used for genotyping during the subsequent seasons and for trials in Naivasha during field evaluation under MLN artificial inoculation.

Genotyping and Marker-Trait Association Analysis

In the present study, we used 19 SNP markers linked to two major QTL associated with resistance to MLN on chromosomes 3 and 6 developed by CIMMYT from two association mapping (AM) panels of diverse tropical/subtropical maize lines (Supplementary Table S1). The two markers explained 33.8% of the phenotypic variance for MLN resistance in the two panels (Gowda et al., 2015). These SNPs were used to design 21 KASP primers for resistance to MLN at BecA-ILRI Hub Laboratory, with three primers for each SNP; two allele-specific forward primers (5'-3') and one common reverse primer (3'-5'; Supplementary Appendix 2).

The 20 parent lines were planted in tray pots filled with sterile soil in a screen house in July 2017. Two seeds were sown per pot with four replicates and each pot was considered

an entry so that a total of 80 entries were generated. No fertilizer was applied and soil moisture was maintained using irrigation. Leaf samples were collected from each entry 14 days after seedling emergence. Trays were carried aseptically onto the bench in the laboratory to avoid contamination. About 8–10 leaf disks per plant were collected from tips of the youngest leaves of each entry using a leaf puncher and placed in a 1.2 ml Eppendorf tube (F and S Scientific Ltd., Kenya) arranged in a 96-well plate placed in an ice bucket. At least 94 wells were filled with samples, while two more were filled with no treatment control (NTC) using ddH₂O.

DNA extraction procedures were based on Cetyl Trimethyl Ammonium Bromide (CTAB) protocols developed by BecA-ILRI Hub Laboratory with some modifications. After sampling, leaf samples were frozen in liquid nitrogen for 2–5 min and ground into fine powder using Geno/Grinder (SPEXSamplePrep^(R) 2000, 2 Dalston Gardens Stanmore, HA7 1BQ, United Kingdom). Genomic DNA was extracted from the fine powder samples following the CTAB protocols. DNA concentration was measured using a spectrophotometer (NanoDrop 8000 Spectrophotometer, Thermo Fisher Scientific, Wilmington, DE, United States) and was adjusted to 50 ng/μl using Nuclease-free water (Patterson et al., 2017). DNA samples of poor quality were discarded and therefore, only samples with high DNA quality were retained and used in this study. Extracted genomic DNA samples were clearly labeled and stored at –20°C until further use (Kusza et al., 2018). The 21 primers generated above were screened for polymorphism to MLN resistance using the 20 parent lines (one resistant donor and 19 susceptible lines) described earlier. Genotypic analyses were conducted using KASP platform established in BecA-ILRI Hub Laboratory in Nairobi, Kenya.

Kompetitive allele specific PCR assays refer to a combination of three SNP-specific primers (two forward and one common reverse), while master mix contained FRET cassettes, free nucleotides, and enzyme components. These materials are required for running the PCR therefore, they should be secured before initiating any PCR process. Correct combination of assays and master mix is vital for obtaining good PCR output and KASP results.

In the present study, both SNP specific KASP Assays and KASP master mix (2xKASP) were ordered from LGC Genomics (LGC Group, Queens Road, Teddington, Middlesex, TW11 0LY, United Kingdom; see Footnote 1). Each assay was supplied in a single 2D barcoded tube with the allele-specific forward primers differing in their tail sequences: allele-1 tail was labeled with fluorescein amide (FAM) oligo-sequence; allele-2 tail contained Hexachloro-fluorescein (HEX) oligo-sequence (Rasheed et al., 2016; Patil et al., 2017; Patterson et al., 2017). KASP master mix was composed of universal FRET cassette dyes (FAM and HEX), ROX™ passive reference dye, KASPTaq™ DNA polymerase, free nucleotides, and MgCl₂ in an optimized buffer solution. On arrival, the assays and the mix were stored at -20°C until further use. The SNP-specific KASP assays (primers) and universal KASP master mix (2xPCR) obtained above were used to constitute KASP reaction mix. Universal KASP master mix was readily obtained from the supplier as described earlier. The reaction mix was constituted in a 100 µl volume as follows:

| | | |
|----------------------------|---|--------|
| FAM (forward primer) | = | 12 µl |
| HEX (forward primer) | = | 12 µl |
| Common primer | = | 30 µl |
| Water (ddH ₂ O) | = | 46 µl |
| Total | = | 100 µl |

Volume of each of KASP master mix and Assay mix required for KASP reaction was aliquoted into new 1.5 ml Eppendorf tube using a pipette as follow:

| | | |
|-----------------|---|-------------------|
| KASP master mix | = | 2.5 µl × N × 1.5 |
| Assay mix | = | 0.07 µl × N × 1.5 |

where N = number of wells to be filled per reaction, 1.5 = error factor. The KASP master mix and Assay mix were then combined into a common volume to constitute a KASP genotyping reaction mix (cocktail). A volume of KASP genotyping reaction mix was aliquoted in each reaction well of a four-quadrant 384-well plate (one quadrant = 96 wells) as below:

| | | |
|----------------------------------|---|--------|
| KASP reaction mix | = | 2.5 µl |
| DNA sample | = | 2.5 µl |
| Total KASP reaction mix per well | = | 5.0 µl |

Assays were put in wells in a 96-well plate where two last wells were stamped with sterile water as NTCs. KASP reaction plate was sealed with optically clear seal and centrifuged at 3,500 rpm for 10–15 s. The template was then amplified using KASP thermal cycling reaction in a FRET-capable plate reader (qPCR) instrument (GeneAmp® PCR System 9700, Roche Molecular Systems, Inc., United States). Unlike other traditional thermal cyclers where three thermal regimes are required, only two temperature steps were used for KASP thermal cycling. In the first step, KASP activation and DNA denaturation was

TABLE 2 | List of selected BC₃F₂ lines including resistant and susceptible parents evaluated for two seasons under artificial inoculations for resistance to MLN in Naivasha in 2018.

| SN | Pedigree | SN | Pedigree |
|----|--|----|------------------------------------|
| 1 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 32 | (CML511 * 2/KS23-6):B-1154 > 1,037 |
| 2 | (CKDHL0186 * 2/KS23-6):B-1019 > 1,033 | 33 | (CML511 * 2/KS23-6):B-1154 > 1,037 |
| 3 | (CKDHL0186 * 2/KS23-6):B-1019 > 1,033 | 34 | (CML511 * 2/KS23-6):B-1154 > 1,037 |
| 4 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,040 | 35 | (CML511 * 2/KS23-6):B-1083 > 1,008 |
| 5 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,040 | 36 | (CML511 * 2/KS23-6):B-1154 > 1,037 |
| 6 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,040 | 37 | (CML511 * 2/KS23-6):B-1083 > 1,008 |
| 7 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 38 | (CML511 * 2/KS23-6):B-1083 > 1,008 |
| 8 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 39 | (CML511 * 2/KS23-6):B-1154 > 1,037 |
| 9 | CKDHL0106 | 40 | CML511 |
| 10 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 41 | (CML511 * 2/KS23-6):B-1154 > 1,037 |
| 11 | CML444 | 42 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 12 | (CML444 * 2/KS23-6):B-1118 > 1,008 | 43 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 13 | (CML444 * 2/KS23-6):B-1118 > 1,008 | 44 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 14 | (CML444 * 2/KS23-6):B-1118 > 1,008 | 45 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 15 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 46 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 16 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 47 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 17 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 48 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 19 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 49 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 20 | (CML511 * 2/KS23-6):B-1154 > 1,041 | 50 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 21 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 51 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 22 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 52 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 23 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 53 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 24 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 54 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 25 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 55 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 26 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 56 | KS23-6 |
| 27 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 57 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 28 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 58 | CML547 |
| 29 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 59 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 30 | KS23-6 | 60 | (CML547 * 2/KS23-6):B-1028 > 1,008 |
| 31 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 61 | (CML547 * 2/KS23-6):B-1092 > 1,019 |
| 31 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 62 | (CML547 * 2/KS23-6):B-1028 > 1,008 |

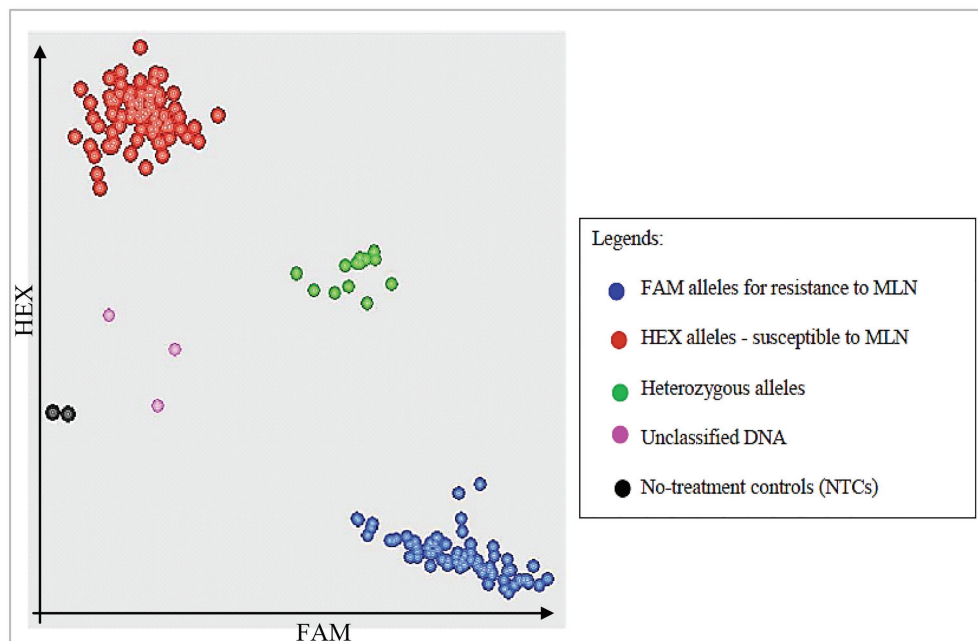


FIGURE 1 | Schematic illustration of kompetitive allele specific PCR (KASP) cluster plots showing segregation of bi-parental backcross (BC) populations for alleles for resistance and susceptibility to maize lethal necrosis (MLN). Fluorescein amide (FAM) alleles for resistance to MLN are clustered on x-axis, while susceptible Hexachloro-flourescein (HEX) alleles are on y-axis.

completed in one cycle and at higher temperature (94°C) in 15 min. Step 2 involved two cycle regimes: Annealing and elongation completed in 10 cycles and at lower temperatures of 61–55°C in 60 s; and the last step required 26 cycles where DNA denaturation occurred at 94°C in 20 s; while annealing and elongation occurred at 55°C in 60 s. Sample amplifications were performed for 30, 35, and 40 cycles. Running the PCR for more than two cycles was necessary to provide an opportunity to select a cycle with the best and clearest clustering of samples for further genotyping (Rasheed et al., 2016; Patil et al., 2017). After KASP reactions were complete, plates were read using fluorescence plate reader BMG FLUOstar Omega software (LGC, Queens Road, Teddington, Middlesex, TW11 0LY, United Kingdom). Data was then displayed as cluster plots where FAM values were plotted on the x-axis, HEX values plotted on y-axis and heterozygous values clustered on the diagonal (**Figure 1**). KASP reactions with NTCs were plotted at the origin (represented by black dots) since they did not generate any fluorescence. Each data point on the cluster plot represented the fluorescence signal of individual DNA samples. Based on the plate readings, two SNP markers (S3_146250249 and S3_146363360) associated with major QTL for resistance to MLN (**Table 3**) showed polymorphisms in the parents, hence were used for genotyping of the backcross populations.

Seeds of BC₃F₂ progeny obtained from the nine populations were planted in a nursery in Kiboko in July 2017. Leaf samples were collected from 250 to 300 healthy individual BC₃F₂ plants per population 15–20 days after seedling emergence. The leaf sampling techniques followed the procedures adopted by

TABLE 3 | Segregation of nine recurrent parents and a donor parent genotyped for resistance to MLN using two SNP markers.

| SN | Parent | SNP1: Chr3_146250249 (T/G) | Chr3_146363360 (C/T) |
|----|----------------------------------|----------------------------------|-------------------------|
| 1 | CML548 | G:G | T:T |
| 2 | CKDHL0106 | G:G | T:T |
| 3 | CML444 | G:G | T:T |
| 4 | CML511 | G:G | T:T |
| 5 | CML547 | G:G | T:T |
| 6 | CML566 | G:G | T:T |
| 7 | CML570 | G:G | T:T |
| 8 | CML539 | G:G | T:T |
| 9 | CKDHL0186 | G:G | T:T |
| 10 | KS23-6 (MLN resistance donor) | T:T | C:C |

T = favorable allele for SNP1 and C = favorable allele for SNP2.

BecA-ILRI Hub Laboratory for field sample collection. Each 96-well plate was labeled and only 94 wells were filled with samples, while two wells were the NTCs as described earlier. Tubes were securely closed with perforated strip caps. To enhance moisture reduction and drying of the leaf samples, a 50 g bag of silica gel (Grade 4) obtained from BecA-ILRI Hub Laboratory was put on top of each 8 × 12 strip tubes in a plate, covered and securely tied with a rubber band, then packaged in a zip-tight polythene bag. Samples were transported to BecA-ILRI Laboratory in Nairobi within 12–24 h and stored at room temperature on the bench until DNA extraction was initiated. Genomic DNA extraction and genotyping followed

TABLE 4 | List of six polymorphic KAPS primers validated for resistance to MLN using 20 parent lines including a resistant donor parent.

| Assay code | Primer name | Primer sequence | Remarks |
|------------|------------------|---|---|
| B0051_Fam | S3_44062810_Fam | gaaggtgaccaagttcatgctATCCGCCTTATTGCCGGg | Polymorphic, linked to QTL with minor effects |
| B0051_HEX | S3_44062810_HEX | gaaggtcggagtgcaacggattATCCGCCTTATTGCCGGa | |
| B0051_COm | S3_44062810_COm | AGGATTAACGACGGGAAGGT | |
| B0054_Fam | S3_146966676_Fam | gaaggtgaccaagttcatgctGTCCTGCTGCTGGAGCGt | Polymorphic, linked to QTL with minor effects |
| B0054_HEX | S3_146966676_HEX | gaaggtcggagtgcaacggattGTCCTGCTGCTGGAGCGc | |
| B0054_COm | S3_146966676_COm | GTAGCGCTCCCGGATGAT | |
| B0056_Fam | S3_146250249_Fam | gaaggtgaccaagttcatgctTACCCATCCGCCTGCTt | Polymorphic, linked to QTL with major effects |
| B0056_HEX | S3_146250249_HEX | gaaggtcggagtgcaacggattTACCCATCCGCCTGCTg | |
| B0056_COm | S3_146250249_COm | CACCTGGCAGCGAGAGAAG | |
| B0057_Fam | S3_146363360_Fam | gaaggtgaccaagttcatgctACCAGGACAGGTATCTAACGCc | Polymorphic, linked to QTL with major effects |
| B0057_HEX | S3_146363360_HEX | gaaggtcggagtgcaacggattACCAGGACAGGTATCTAACGCt | |
| B0057_COm | S3_146363360_COm | CGTACCAGGTCTGAGCACAA | |
| B0060_Fam | S6_21007530_Fam | gaaggtgaccaagttcatgctGCAAAAATCACAGCCGATCg | Polymorphic, linked to QTL with minor effects |
| B0060_HEX | S6_21007530_HEX | gaaggtcggagtgcaacggattGCAAAAATCACAGCCGATCa | |
| B0060_COm | S6_21007530_COm | CCGGGCCTAAAGCCTAATAC | |
| B0065_Fam | S6_157568432_Fam | gaaggtgaccaagttcatgctGCATAGAAATAAAATGAGACAAGGg | Polymorphic, linked to QTL with minor effects |
| B0065_HEX | S6_157568432_HEX | gaaggtcggagtgcaacggattGCATAGAAATAAAATGAGACAAGGt | |
| B0065_COm | S6_157568432_COm | ATCCATGTTGTCCCTCCGTA | |

the same protocol as described for parental screening. The two SNP markers identified above were used for the genotyping using Omega software. Fifty-seven BC₃F₂ progeny carrying the two markers were identified through the genotyping.

Data from Omega software were imported into KlusterCaller software (LGC Genomics, Queens Road, Teddington, Middlesex, TW11 0LY, United Kingdom) and cluster plots were normalized using ROX (passive reference dye) then called into X:X, X:Y, and Y:Y alleles depending on the corresponding genotype. Results were then exported onto Excel 2016 version, following 96-well plate format and calls were converted into specific alleles where X:X represented homozygous alleles for FAM, Y:Y represented homozygous alleles for HEX, and X:Y represented segregating (heterozygous) alleles (FAM/HEX; Graves et al., 2016; Kusza et al., 2018). The KASP analysis revealed that six of the SNP primers were polymorphic and could clearly discriminate between resistance alleles of donor parent (KS23-6) and the susceptibility alleles of the recipient parents (Table 4). The remaining 15 SNPs were monomorphic and could not differentiate between the resistant and susceptible parents. Previous reports confirm that two out of the six polymorphic SNPs are linked to major QTLs for resistance to MLN, while the remaining four markers are also for resistance but with minor effects (Gowda et al., 2015, 2018). As a result, only the two major SNPs (S3_146250249 and S3_146363360) were retained and used as markers for genotyping of the BC₃F₂ populations.

Phenotypic Evaluation of Backcross Populations

The 57 BC₃F₂ lines selected from the molecular analysis for this study, including four checks (one MLN resistant and three MLN susceptible) and their parents were evaluated for two seasons in 2018 to validate the effects of the two markers for resistance to MLN under MLN artificial inoculation in the field. The two markers, located approximately 0.62 cM (113111 nt)

apart on chromosome 3, were putatively identified in different populations using genome-wide association studies (Gowda et al., 2015). Experiments were conducted in Naivasha, Kenya (36°26'E; 0°43'S; 1896 masl; 677 mm rainfall; and 24.9°C; Ziyomo and Bernardo, 2013; Odiyo et al., 2014). To reduce soil borne diseases and pest infections, seeds were treated with Apron Star WS seed treatment chemical at the rate of 20 g/kg of seed. Recommended fertilizer rates were adopted and applied in two separate regimes (Izge and Dugje, 2011). The trial was laid out in an alpha lattice design with two replications, and a one-row plot of 3.0 m, with spacing of 0.75 m between rows and 0.3 m between plants. Two seeds were planted per hill and thinned to one plant per hill 3 weeks post emergence, resulting in a total of 10 plants per row. Standard agronomic practices were maintained (Gowda et al., 2015; Mahuku et al., 2015).

Artificial MLN Inoculation

Infected leaf samples collected from the field were cut into small pieces and ground using a mortar and pestle in a grinding buffer of 1:10 dilution ratio (10 ml potassium-phosphate, pH 7.0) as described by Gowda et al. (2015) and Mahuku et al. (2015). The resulting sap extract was centrifuged for 2 min at 12,000 rpm. Celite powder was added to the decanted sap extract at the rate of 0.02 g/ml. A susceptible hybrid was inoculated by rubbing sap extract onto the leaves at the two leaf stage and infected maize plants grown in separate, sealed greenhouses that were maintained for each of SCMV and MCMV inoculum production. Three weeks before inoculation of the experimental materials, ELISA test was conducted on random samples of leaves from the plants infected with SCMV and MCMV, respectively, to confirm presence and purity of the inoculum (Gowda et al., 2018). Separate extracts from the SCMV and MCMV infected plants were prepared at the ratio of one part of leaf sample: 20 parts of phosphate buffer. The two extracts were then mixed to form MLN inoculum at the ratio of four parts of SCMV: one part MCMV (weight/weight;

Gowda et al., 2015). In order to keep uniform disease pressure, plants were inoculated using a motorized, backpack mist blower (Solo 423 mist Blower, 12 L capacity) with an open nozzle (2-in diameter) delivering inoculum spray at a pressure of 10 kg/cm² (Gowda et al., 2015). Two inoculations were applied at 4th and 5th week after planting (Gowda et al., 2018). Spreader rows of susceptible maize hybrid (H614) were also planted as border rows along the experiment to enhance disease spread and intensity (Tivoli et al., 2006). Nitrogen (Urea) and Phosphorus (DAP) fertilizers were applied as described by Makumbi et al. (2015). Drip irrigation was used to provide moisture and all other agronomic practices relating to maize production were followed according to CIMMYT procedures for field practices.

A quantitative scale of 1–9 introduced by Reddy and Singh (1984) was used for recording data on MLN severity, where 1 = resistant (no symptoms); 2 = resistant to moderately resistant (isolated plants with very few lesions in the lower canopy); 3 = moderately resistant (1–5 leaves with symptoms in the lower canopy); 4 = moderately resistant to moderate (most or all plants with one or more leaves affected in the lower canopy); 5 = moderate (most or all plants with many leaves affected on plant, few leaves affected in the mid canopy); 6 = moderate to moderately susceptible (numerous lesions on most leaves in the mid canopy, limited defoliation in lower canopy); 7 = moderately susceptible (same as six, but limited defoliation in mid canopy and severe defoliation in lower canopy); 8 = moderately susceptible to susceptible (severe defoliation in mid canopy and limited defoliation in upper canopy); and 9 = susceptible (complete plant necrosis; Ngugi et al., 2002; Meyer and Pataky, 2010). The scale of 1–9 was considered more convenient (in terms of recording and time) compared to 1–5 because: 1 = 1; 1.5 = 2; 2 = 3; 2.5 = 4; 3 = 5; 3.5 = 6; 4 = 7; 4.5 = 8; and 5 = 9. Disease severity was recorded four times, beginning 21 days from the date of first inoculation (Mahuku et al., 2015; Mezzalama, 2015; Gowda et al., 2018).

Disease severity data were first tested for independence, normal distribution and constant variance (GenStat ver 12.0). ANOVA was performed using the restricted maximum likelihood (REML) model established in SAS 9.4 (SAS institute Inc, 2016), based on lattice incomplete block analysis as follows:

$$Y_{ijk} = \mu + g_i + r_j + b_{kj} + e_{ijk}$$

where Y_{ijk} is the disease severity of the i th genotype in the j th replication of the k th incomplete block, μ is the population mean, g_i is the genetic effect of the i th genotype, r_j is the effect of the j th replication, b_{kj} is the effect of the k th incomplete block in the j th replication, and e_{ijk} is the error term. Genotype was considered fixed, while season, replication, and block within replication were considered random. Best linear unbiased prediction estimates (BLUP) for the populations were generated and used for further QTL analysis in the present study.

Marker-Trait Association Analysis

The BLUP estimate for phenotypic data was generated for each genotype as described above. Co-segregation of loci with phenotype was detected by comparing allele type with the phenotype.

Positive co-segregation was declared when a genotype showed resistance allele and MLN score of below 4.0 (in the scale of 1–9). Relationship between allele and phenotype was confirmed by splitting the BC₃F₂ progeny into resistant and susceptible groups and means of the two groups compared using a t -test. Further, significance of differences between variances of means was determined using F -tests.

RESULTS

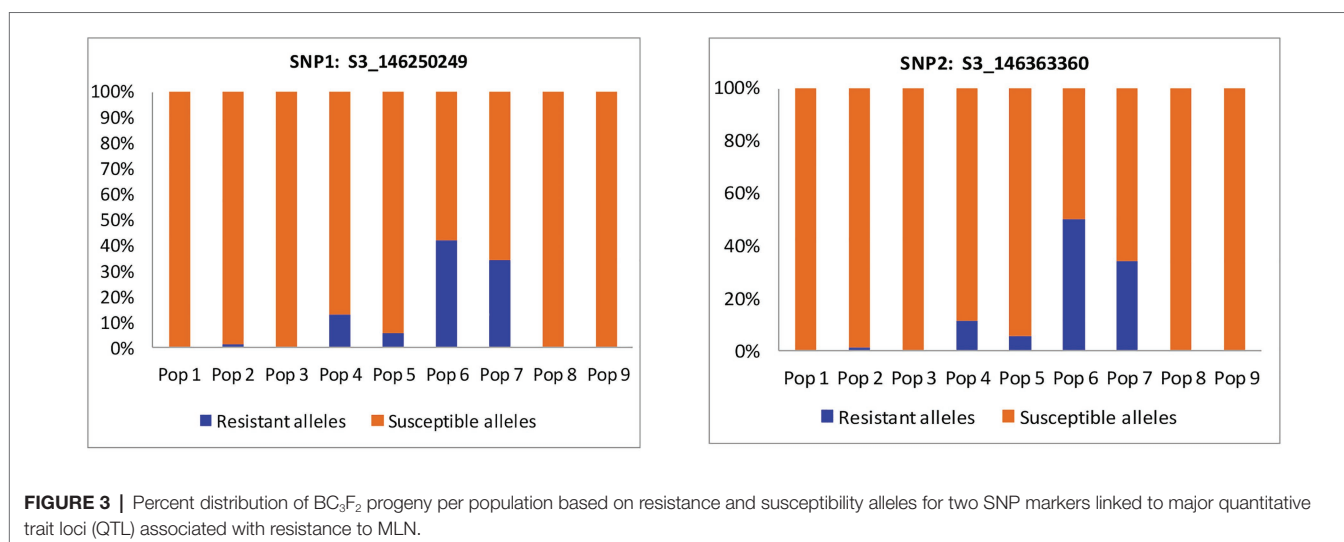
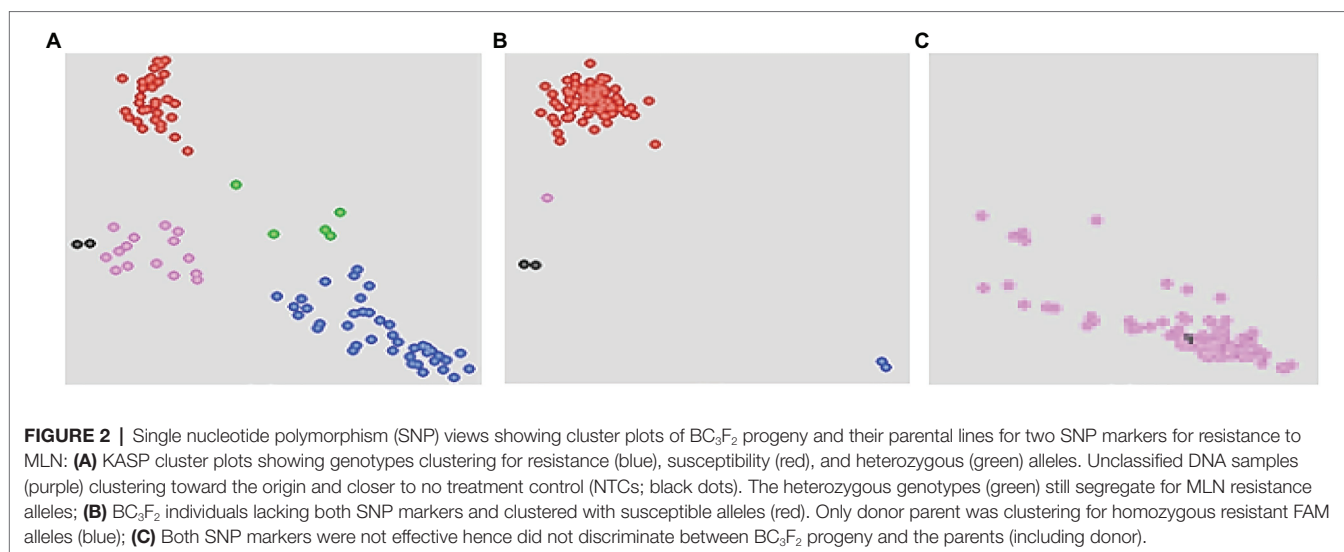
Screening of Parental Lines

A total of 21 KASP primers designed from 19 SNP markers, originally developed by CIMMYT Global Maize Program, were tested for polymorphism to MLN resistance using 20 parental lines that were involved in development of bi-parental backcross populations. Six KASP markers showed polymorphism for resistance to MLN among 9 of the 19 bi-parental backcross populations used. As a result, the remaining 11 populations were eliminated. Two of the six polymorphic KASP markers (S3_146250249 and S3_146363360) were previously reported to be linked to major QTL associated with resistance to MLN (Gowda et al., 2015). Therefore, they were retained for this study. The preliminary KASP analysis revealed that recurrent parents of the nine selected populations were fixed for susceptibility alleles, while the donor parent was homozygous for resistance alleles for both markers (Table 3).

Genotyping of BC₃F₂ Populations for Resistance to MLN

Selected high quality DNA samples representing 957 BC₃F₂ lines selected from nine bi-parental populations and their 10 parental lines were genotyped for resistance to MLN using KASP genotyping platform at Beca-ILRI Hub Laboratory, Nairobi. The two polymorphic KASP markers mentioned earlier linked to major QTL for resistance to MLN were used (Semagn et al., 2014; Rasheed et al., 2016; Kusza et al., 2018). KASP results showed clustering of the genotypes based on the two KASP markers. Some BC₃F₂ and all recurrent parents clustered with susceptible homozygous HEX alleles on y-axis. The donor parent and some BC₃F₂ lines clustered with the resistant homozygous FAM alleles on x-axis. A few BC₃F₂ lines clustered for heterozygous alleles on the diagonal. However, the two markers failed to discriminate between some BC₃F₂ including parents (Figure 2).

Cluster plot results indicated that 57 BC₃F₂ individuals were segregating for resistance to MLN. A total of 26 BC₃F₂ lines were homozygous (fixed) for the favorable alleles of both KASP1 (S3_146250249) and KASP2 (S3_146363360). The remaining BC₃F₂ lines were fixed for one locus and heterozygous for the other (Supplementary Appendix 3). The selected 57 BC₃F₂ lines were subjected to artificial MLN infection for phenotypic selection under field conditions. Allele distribution for each SNP marker varied among the nine populations (Figure 3). Both markers showed higher percentages of alleles for resistance to MLN among progeny from populations six and seven.



Population 6 contributed most with over 40% of the population showing favorable alleles for resistance to MLN, followed by population 7 with over 30% distribution of BC₃F₂ progeny containing favorable alleles for resistance to MLN for both markers. Populations 1, 2, 3, and 9 contributed the least with less than 20% each for both markers.

Phenotypic Evaluation of BC₃F₂ Populations for Resistance to MLN

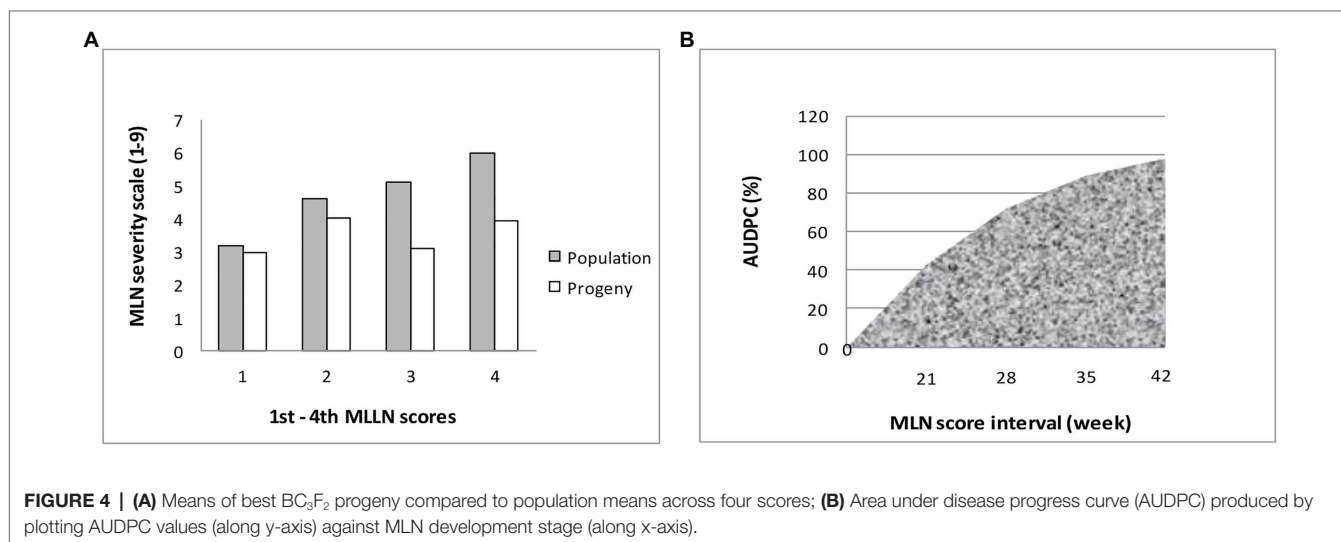
The mean distribution for the populations used in the current study is presented in **Table 5**. The results showed that mean distribution for MLN scores ranged from 3.2 for first score to 6.0 for the last fourth score, with area under disease progress curve (AUDPC) of 133.2, respectively. A total of 19 selected BC₃F₂ lines including MLN resistance donor parent showed MLN severity and AUDPC below the population mean for both early and late scores. BCL02 was the most resistant line with MLN score of 3.1, which was similar to the mean of

the donor parent and it had an even lower AUDPC score of 69.6 compared to the donor parent that showed a score of 81.8. When mean severity (1–9) was plotted against the score interval (in weeks), it was observed that the best performing BC₃F₂ lines had lower MLN mean severity across scores compared to the general population mean (**Figure 4A**). Similarly, the trend of the development of AUDPC showed that both disease severity and AUDPC values increased as MLN severity increased from first to fourth scores (**Figure 4B**). Results obtained from ANOVA are shown in **Table 6**. It was observed that there were variations among the BC₃F₂ lines for response to MLN infections under field conditions. The variability among the genotypes ranged from significant ($p \leq 0.05$) at first MLN severity score (MLN1) to highly significant ($p \leq 0.01$) for the fourth MLN severity score (MLN4). Similarly, the results showed highly significant ($p \leq 0.001$) variability for AUDPC. Broad-sense heritability was detected to be very high and it ranged from of 0.84 to 0.91. Narrow-sense heritability was moderate to high with values ranging from 0.32 to 0.58, respectively.

TABLE 5 | Mean scores and AUDPC for MLN severity for 10 resistant BC₃F₂ lines compared to donor and susceptible parents evaluated under artificial MLN infections in Naivasha for two seasons in 2018.

| SN | Geno. | Pedigree | MLN1 | MLN2 | MLN3 | MLN4 | AUDPC |
|----|------------|--|------|------|------|------|-------|
| 1 | BCL02 | (CKDHL0186 * 2/KS23-6):B-1019 > 1,033 | 2.5 | 2.5 | 2.4 | 3.1 | 69.6 |
| 2 | BCL05 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 2.8 | 3.6 | 2.1 | 3.2 | 69.5 |
| 3 | BCL16 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.8 | 3.6 | 2.5 | 3.2 | 70.3 |
| 4 | BCL03 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 3.0 | 3.9 | 2.3 | 3.3 | 73.4 |
| 5 | BCL18 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.8 | 4.1 | 3.4 | 3.4 | 82.2 |
| 6 | BCL20 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.9 | 4.0 | 2.6 | 3.7 | 76.7 |
| 7 | BCL17 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 3.0 | 3.7 | 2.8 | 3.9 | 77.8 |
| 8 | BCL27 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 3.0 | 4.1 | 3.1 | 3.9 | 81.6 |
| 9 | BCL31 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.9 | 4.3 | 2.9 | 4.0 | 82.8 |
| 10 | BCL33 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.9 | 4.1 | 2.9 | 4.0 | 80.3 |
| 11 | Check | KS23-6 (Donor parent) | 2.9 | 4.5 | 3.0 | 3.1 | 81.8 |
| 12 | Check | CKDHL0106 (Recurrent parent) | 3.3 | 5.0 | 4.8 | 5.7 | 116.6 |
| 13 | Check | CML511 (Recurrent parent) | 3.6 | 5.1 | 5.5 | 6.6 | 134.5 |
| | Mean | | 3.2 | 4.6 | 5.1 | 6.0 | 133.3 |
| | LSD (0.05) | | 0.6 | 1.2 | 0.9 | 1.2 | 22.7 |
| | CV% | | 11.8 | 14.6 | 10.3 | 12.1 | 9.9 |
| | Min mean | | 1.8 | 2.4 | 2.1 | 3.1 | 69.5 |
| | Max mean | | 3.9 | 6.5 | 8.1 | 8.8 | 212.7 |

MLN severity scores taken on a scale of 1–9, where 1–4 is resistant to moderate resistant. MLN1, 1st score of MLN severity taken 21 days from date of first inoculation; MLN2, 2nd score of MLN severity taken 7 days after the first score; MLN3, 3rd score of MLN severity recorded 14 days after the first score; MLN4, 4th score of MLN severity taken 21 days from the first score; AUDPC, area under disease progress curve calculated from the four MLN scores; LSD (0.05), Fisher's protected least significant difference at 5% level; and CV%, coefficient of variability measures in percent.



Validation of Marker Effects on Phenotypic Variations

At least 57 BC₃F₂ lines selected using the two KASP markers linked to major QTL for resistance to MLN were evaluated together with their parents under artificial MLN inoculation in Naivasha during first season of 2018. Phenotypic means and genotypic data were compared and co-segregations of resistance alleles with phenotype were determined (Supplementary Appendix 4). Some selected BC₃F₂ lines with strong allele-phenotype associations for resistance to MLN are shown in Tables 7 and 8. Additionally, 31 BC₃F₂ lines fixed for one of the two resistance loci showed resistant to moderately susceptible reactions to MLN with means ranging from 3.1 to 6.7. However, 20 BC₃F₂ lines fixed for both resistance loci,

showed susceptibility to MLN with mean severity of 7.5–8.8 (data not shown).

Distribution of genotypes based on their responses to MLN infection is shown in Figure 5. The recurrent parents demonstrated moderate to highly susceptible responses to MLN. Comparison between early and late scores showed that early MLN mean severity values for all genotypes were below 4.0 (Figure 5A), however, for later scores, a number of individual plants succumbed to MLN infection with disease scores above 8 (Figure 5B). Out of 57 BC₃F₂ lines genotyped, six were fixed for both resistance loci and showed high resistance to MLN under MLN artificial inoculation in the field, whereas 31 lines fixed for only one of the two loci demonstrated moderate resistance to the disease. Another 20 lines though

fixed for both resistance loci, however, manifested high susceptibility to MLN under artificial MLN inoculation in the field (**Figure 5C**). Mean scores for resistant genotypes was compared to the means of the susceptible group. The resistant category demonstrated lower mean MLN score of 3.9 compared to 7.0 for the susceptible genotypes (**Figure 5D**). Means of the two groups were subjected to *t*-test and the results revealed highly significant differences ($p \leq 0.001$) between means of resistant and susceptible groups of the populations (**Table 9**). Further, differences between variances of the two means were

determined using the *F*-test and the results showed highly significant differences ($p \leq 0.0001$) as shown below (**Table 8**). Consequently, 19 elite BC₃F₂ lines fixed for both or one locus and showing resistant to moderately resistant reaction to MLN infection were identified.

DISCUSSION

Kompetitive allele specific PCR analysis showed clustering of some BC₃F₂ lines with the donor parent. This indicates that the lines could be fixed for favorable alleles of the two KASP markers for resistance to MLN. The high MLN severity observed at fourth score was attributed to an increase in disease severity with time. The genotypes under study revealed two categories in the field based on their responses to MLN incidence. The first category showed low MLN scores. This might imply that the two QTL were stable across different genetic backgrounds. Therefore, materials in this category were fixed for the two loci, hence were able to manifest resistance under field conditions and minimize MLN effects. The second group, though fixed for both loci, showed high susceptibility under field conditions, which could be due to false positive effects. This means the QTL were falsely selected for favorable alleles, while they were carrying susceptibility alleles. The reason could be that the KASP markers separated from the resistance genes during

TABLE 6 | Mean squares and variance components of BC₃F₂ populations evaluated for two seasons for resistance to MLN in Naivasha in 2018.

| Trait | Mean squares | F-value | F-prob. | H ² | σ^2_G | σ^2_E |
|-------|--------------|---------|---------|----------------|--------------|--------------|
| MLN1 | 0.4 | 2.25 | 0.024 | 0.84 | 0.92 | 0.17 |
| MLN2 | 1.4 | 2.7 | 0.008 | 0.87 | 3.23 | 0.50 |
| MLN3 | 1.6 | 4.74 | <0.001 | 0.91 | 3.49 | 0.33 |
| MLN4 | 2.1 | 3.09 | <0.001 | 0.88 | 4.84 | 0.67 |
| AUDPC | 834.8 | 3.86 | <0.001 | 0.90 | 1886.10 | 216.50 |

MLN1, 1st score of MLN severity taken 21 days from date of first inoculation; MLN2, 2nd score of MLN severity taken 7 days after the first score; MLN3, 3rd score of MLN severity recorded 14 days after the first score; MLN4, 4th score of MLN severity taken 21 days from the first score; AUDPC, area under disease progress curve calculated from the four MLN scores; H², broad sense heritability; σ^2_G , genotypic variance; and σ^2_E , error variance; the Degree of freedom = 61 for each trait.

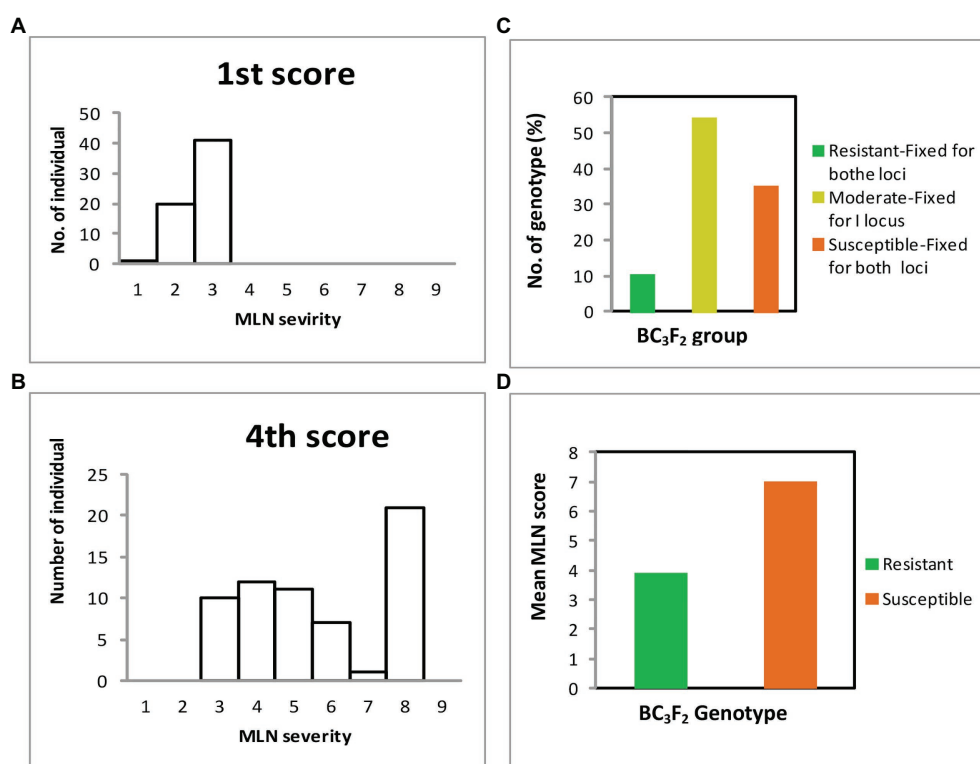


FIGURE 5 | Responses of BC₃F₂ lines to MLN infections in the field recorded at scale of 1–9: **(A)** mean distribution of MLN severity for first score; **(B)** mean distribution of MLN severity fourth score; **(C)** mean distribution of MLN scores for resistant, moderate, and susceptible genotypes; and **(D)** mean distribution of MLN scores for resistant and susceptible groups of genotypes.

TABLE 7 | Some selected BC₃F₂ lines and a donor parent showing strong co-segregation of resistant alleles with phenotypic MLN scores under field infections.

| S/N | Genotype | SNP1 (T/G) | SNP2 (C/T) | MLN1 | MLN2 | MLN3 | MLN4 |
|-----|----------------------|------------|------------|------|------|------|------|
| 1 | KS23-6 (door parent) | T:T | C:C | 2.9 | 4.5 | 3.0 | 3.1 |
| 2 | BCL5 | T:T | C:C | 2.8 | 3.6 | 2.1 | 3.2 |
| 3 | Check5 | T:T | C:C | 1.8 | 2.4 | 3.0 | 3.2 |
| 4 | BCL3 | T:T | C:C | 3.0 | 3.9 | 2.3 | 3.3 |
| 5 | BCL10 | T:T | C:C | 3.0 | 4.4 | 3.6 | 4.2 |
| 6 | BCL11 | T:T | C:C | 2.8 | 4.1 | 3.1 | 4.2 |

T:T and C:C indicate MLN resistant loci for SNP1 and SNP2.

TABLE 8 | List of 19 MLN resistant BC₃F₂ lines selected for testcross development.

| SN | Pedigree | MLN1 | MLN2 | MLN3 | MLN4 | AUDPC |
|----------|--|------|------|------|------|--------|
| 1 | (CKDHL0186 * 2/KS23-6):B-1019 > 1,033 | 2.1 | 2.1 | 2.5 | 3.0 | 69.6 |
| 2 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 2.1 | 2.1 | 2.0 | 3.0 | 69.5 |
| 3 | (CKDHL0106 * 2/KS523-5):B-1110 > 1,016 | 2.1 | 2.1 | 2.0 | 3.0 | 70.3 |
| 4 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.1 | 2.5 | 3.0 | 73.4 |
| 5 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.7 | 3.0 | 3.0 | 82.2 |
| 6 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.2 | 2.5 | 3.5 | 76.7 |
| 7 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.7 | 2.5 | 3.5 | 77.8 |
| 8 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.1 | 2.2 | 2.5 | 3.5 | 81.6 |
| 9 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.1 | 2.7 | 2.5 | 3.5 | 82.8 |
| 10 | (CML444 * 2/KS23-6):B-1118 > 1,008 | 2.6 | 3.6 | 3.5 | 4.0 | 80.3 |
| 11 | (CML444 * 2/KS23-6):B-1118 > 1,008 | 2.1 | 3.1 | 3.0 | 4.0 | 78.9 |
| 12 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.2 | 3.0 | 4.0 | 85.4 |
| 13 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.2 | 3.0 | 4.0 | 81.6 |
| 14 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.2 | 3.0 | 4.0 | 100.1 |
| 15 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.1 | 3.2 | 3.0 | 4.0 | 88.6 |
| 16 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.1 | 3.2 | 3.0 | 4.0 | 81.4 |
| 17 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.1 | 2.6 | 3.0 | 4.0 | 85.2 |
| 18 | (CML511 * 2/KS23-6):B-1154 > 1,037 | 2.1 | 2.7 | 3.0 | 4.0 | 86.5 |
| 19 | (CML511 * 2/KS23-6):B-1083 > 1,008 | 2.1 | 2.2 | 3.0 | 4.5 | 85.2 |
| Mean | | 3.17 | 4.60 | 5.09 | 6.03 | 133.28 |
| LSD (5%) | | 0.40 | 0.74 | 0.55 | 0.73 | 16.87 |
| CV % | | 6.80 | 8.17 | 5.98 | 7.18 | 5.81 |
| H^2 | | 0.94 | 0.95 | 0.99 | 0.98 | 0.97 |
| h^2 | | 0.32 | 0.39 | 0.45 | 0.52 | 0.58 |

The mean MLN scores and AUDPC were obtained under artificial MLN infections in Naivasha during 2018 cropping season. MLN1, MLN2, MLN3, and MLN4 were 1st, 2nd, 3rd, and 4th MLN severity scores.

meiosis resulting in the markers being present but the genes are not. This outcome highlights the critical importance of confirming resistance of MLN under field conditions when molecular markers are used to select for resistance. The significant ($p \leq 0.01$) variability observed among the genotypes for resistance to MLN, in this study implies that the genotypes responded differently to MLN infection. Further, statistical results showed significant AUDPC and with increased disease development for MLN, with maximum at the fourth score. Development of MLN within plant systems is rapid and is supported by viral movement and replication proteins produced by the pathogens (Mbega et al., 2016; Xia et al., 2016). Therefore, MLN quickly developed such that by the 42nd day (4th scores) after first inoculation, the disease had advanced and colonized most parts of the plant systems leading to expanded AUDPC.

Narrow sense heritability estimates for MLN scores were moderate to high indicating that resistance to MLN was mostly conditioned by additive gene action as opposed to non-additive inheritance. Additionally, the high heritability indicated that

genotype played major roles in influencing the variability in MLN resistance among the individuals compared to the environment in which the experiment was conducted. In summary, the moderate to high narrow-sense heritability estimates imply the ease of transfer of the target trait from parent to offspring. High heritability estimates for disease resistance have been reported in maize (Gowda et al., 2015; Sukruth et al., 2015; Cao et al., 2017). Gowda et al. (2018) reported moderate to high heritability estimates of 0.34–0.89 for early and late MLN scores. Beyene et al. (2017) observed broad sense heritability of 69–73% for MLN resistance. The moderate to high narrow-sense heritability estimates observed in the present study also indicates that genotypes contributed significantly to the phenotypic variation observed. Therefore, identification of these lines for resistance to MLN is possible through field evaluation of the genotypes similar to findings of Pereira et al. (2015).

The present study validated two KASP markers (S3_146250249 and S3_146363360) on chromosome 3, which have been reported to harbor a hot spot region for various genes responsible for

TABLE 9 | Summary of *t*-test of means and *F*-test statistics for the significance of the difference between variances of means for resistance to MLN between resistant and susceptible groups of BC₃F₂ populations.

| Df | 60 |
|---|--------|
| Mean _(Res) | 3.9 |
| Mean _(Sus) | 7.2 |
| Mean _(Res) – mean _(Sus) | 3.35 |
| <i>t</i> -value | –10.6 |
| <i>F</i> -value | 7.4 |
| Prob. | 0.0001 |

Df, degrees of freedom; $p \leq 0.0001$ = means and variances of means are highly significantly different.

resistance to diseases of economic importance in maize (Gowda et al., 2015; Lohithaswa et al., 2015). *T*-test analysis revealed that means of resistant and susceptible groups were highly significant ($p \leq 0.0001$) meaning that resistant and susceptible genotypes performed differently under MLN infections. Similarly, *F*-tests showed that differences between variances of the means were highly significant ($p \leq 0.0001$). The findings implied that phenotypic resistance demonstrated by genotypes was highly related to the favorable alleles associated with major QTL for resistance to MLN. Consequently, 26 BC₃F₂ lines containing both KASP markers demonstrated resistance to MLN infections, suggesting that the two QTL were associated with phenotypic resistance in those populations. Tanweer et al. (2015) used marker assisted backcrossing for introgression of two blast resistance genes (*Pi-b* and *Pi-kh*) into a locally adapted rice line. Evaluation for blast resistance under field conditions revealed that the improved lines had higher resistance against pathotype P7.2. However, a QTL may be transferred into recipient background yet its effect may not show due to interactions (epitasis and linkage) with other genes in the new backgrounds (Hospital, 2005; Collard and Mackill, 2008). In the current study, 20 BC₃F₂ lines were fixed for both KASP markers but were susceptible to MLN under field conditions, which could mainly be due to weak associations (in terms of genetic distance) between the KASP markers and target gene. This could indicate the possibility of false-positive detection of QTL. The susceptible genotypes could be because of separation of these markers with the gene. This probably indicates weak linkage between the markers and the resistance genes. We recommend the identification of more closely linked loci between these markers and the resistance genes.

CONCLUSION

The current study confirmed presence of two KASP markers (S3_146250249 and S3_146363360) with major effects for resistance to MLN under field conditions. Both QTL are located on chromosome 3 at a distance of 113,111 nucleotides apart. These two QTL were reproducible under different genetic and environmental conditions. The validation study confirmed that 19 superior BC₃F₂ lines were fixed for favorable alleles of the two QTL, and showed higher levels of resistance to MLN under artificial field inoculations. These elite BC₃F₂ lines represent

useful parents for developing maize hybrids with resistance to MLN. Furthermore, the validated QTL can be utilized to speed up marker assisted breeding for resistance to MLN. The study identified 20 lines fixed for two KASP markers for resistance to MLN but with susceptible reaction under artificial MLN inoculations suggesting weak marker-gene linkage. We recommend the identification of additional loci between these markers and the resistance genes to strengthen the linkage. The results highlight the importance of confirmation of resistance under field conditions when molecular markers are used for selection.

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

LA conducted experiments, collected data, conducted data analysis, and drafted the manuscript. NY and OS coordinated lab work at BecA, supported molecular analyses, and revised the manuscript. MJ and LS supervised field experiment, coordinated the work in CIMMYT, and reviewed the manuscript. PT, ED, BI, and PM-D supervised the experiment and revised the manuscript. MG, MO, and PB revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was part of the PhD thesis of the LA, funded by Intra-ACP CSAA Mobility Schemes at West Africa Center for Crop Improvement (WACCI), University of Ghana, Legon. The thesis was partially supported through the Africa Centers of Excellence for Development Impact (ACE Impact) project. The genotyping and molecular analysis for this study was funded by BecA-ILRI Hub through Africa Biosciences Challenge Fund Fellowship.

ACKNOWLEDGMENTS

We would like to thank CIMMYT for providing technical support and access to MLN facilities for conducting the experiments for this study. We very much appreciate the contribution from Mrs. Lucy Muthui of SEGOLILAB (BecA-ILRI Hub) for her support in the molecular analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.649308/full#supplementary-material>

REFERENCES

- Acquaah, G. (2012). *Principles of Plant Genetics and Breeding. 2nd Edn.* West Sussex, UK: John Wiley and Sons, Ltd.
- Awata, L. A. O., Beyene, Y., Gowda, M., Jumbo, M. B., Tongoona, P., Danquah, E., et al. (2019). Genetic analysis of QTL for resistance to maize lethal necrosis in multiple mapping populations. *Genes* 11:32. doi: 10.3390/genes11010032
- Awata, L. A. O., Tongoona, P., Danquah, E., Efe, B. E., and Marchelo-dragga, P. W. (2018). Common mating designs in agricultural research and their reliability in estimation of genetic parameters. *J. Agric. Vet. Sci.* 11, 16–36.
- Beyene, Y., Gowda, M., Mugo, S., Suresh, L. M., Olsen, M., Oikeh, S. O., et al. (2017). Genetic analysis of tropical maize inbred lines for resistance to maize lethal necrosis disease. *Euphytica* 213:224. doi: 10.1007/s10681-017-2012-3
- Bouchez, A., Hospital, F., Causse, M., Gallais, A., and Charcosset, A. (2002). Marker-assisted introgression of favorable alleles at quantitative trait loci between maize elite lines. *Genetics* 162, 1945–1959.
- Brandes, E. W. (1920). Artificial and insect transmission of sugarcane mosaic. *J. Agric. Res.* 19, 131–138.
- Brown, J., and Caligari, P. D. S. (2008). *An Introduction to Plant Breeding.* Oxford: Blackwell Publishing.
- Cao, S., Alexander, L., Yibing, Y., Yongsheng, W., Ao, Z., Jiafa, C., et al. (2017). Genome-wide analysis of tar spot complex resistance in maize using genotyping-by-sequencing SNPs and whole-genome prediction. *Plant Genome* 10, 1–14. doi: 10.3835/plantgenome2016.10.0099
- Collard, B. C. Y., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170
- Feng, F., Wang, Q., Liang, C., Yang, R., and Li, X. (2015). Enhancement of tocopherols in sweet corn by marker-assisted backcrossing of ZmVTE4. *Euphytica* 206, 513–521. doi: 10.1007/s10681-015-1519-8
- Gowda, M., Beyene, Y., Makumbi, D., Semagn, K., Olsen, M. S., Bright, J. M., et al. (2018). Discovery and validation of genomic regions associated with resistance to maize lethal necrosis in four biparental populations. *Mol. Breed.* 38:66. doi: 10.1007/s11032-018-0829-7
- Gowda, M., Das, B., Makumbi, D., Babu, R., Semagn, K., and Prasanna, B. M. (2015). Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theor. Appl. Genet.* 128, 1957–1968. doi: 10.1007/s00122-015-2559-0
- Graves, H., Rayburn, A. L., Gonzalez-Hernandez, J. L., Nah, G., Kim, D. S., and Lee, D. K. (2016). Validating DNA polymorphisms using KASP assay in prairie cordgrass (*Spartina pectinata* link) populations in the U.S. *Front. Plant Sci.* 6:1271. doi: 10.3389/fpls.2015.01271
- Hospital, F. (2003). “Marker-assisted Breeding,” in *Plant Molecular Breeding. 1st Edn.* eds. J. A. Roberts and P. P. N. R. Usherwood (Garsington, UK: Blackwell Publishing Ltd.), 30–59.
- Hospital, F. (2005). Selection in backcross programmes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1503–1511. doi: 10.1098/rstb.2005.1670
- Izge, A. U., and Dugie, I. Y. (2011). Performance of drought tolerant three-way and top cross maize hybrids in Sudan savanna of north eastern Nigeria. *J. Plant Breed. Crop Sci.* 3, 269–275.
- Kusza, S., Csiszter, L. T., Ilie, D. E., Sauer, M., Padeanu, I., and Gavojdian, D. (2018). Kompetitive allele specific PCR (KASP™) genotyping of 48 polymorphisms at different caprine loci in French alpine and saanen goat breeds and their association with milk composition. *PeerJ* 6:e4416. doi: 10.7717/peerj.4416
- Lohithaswa, H. C., Jyothi, K., Sunil Kumar, K. R., and Hittalmani, S. (2015). Identification and introgression of QTLs implicated in resistance to sorghum downy mildew (*Peronosclerospora Sorghi* (Weston and Uppal) C. G. Shaw) in maize through marker-assisted selection. *J. Genet.* 94, 741–748. doi: 10.1007/s12041-015-0590-1
- Lübberstedt, T., Ingvarsdson, C., Melchinger, A. E., Xing, Y., Salomon, R., and Redinbaugh, M. G. (2006). Two chromosome segments confer multiple Potyvirus resistance in maize. *Plant Breed.* 125, 352–356. doi: 10.1111/j.1439-0523.2006.01244.x
- Mahuku, G., Lockhart, B. E., Wanjala, B., Jones, M. W., Kimunye, J. N., Stewart, L. R., et al. (2015). Maize lethal necrosis (MLN), an emerging threat to maize-based food security in sub-saharan Africa. *Phytopathology* 105, 956–965. doi: 10.1094/PHYTO-12-14-0367-FI
- Makumbi, D., Alpha, D., Kanampiu, F., Mugo, S., and Karaya, H. (2015). Agronomic performance and genotype × environment interaction of herbicide-resistant maize varieties in eastern africa. *Crop Sci.* 55, 540–555. doi: 10.2135/cropsci2014.08.0593
- Mbega, E. R., Ndakidemi, P. A., Mamiro, D. P., Mushongi, A. A., Kitege, K. M., and Ndonga, O. A. (2016). Role of potyviruses in synergistic interaction leading to maize lethal necrotic disease on maize. *Int. J. Curr. Microbiol. App. Sci.* 5, 85–96. doi: 10.20546/ijcmas.2016.506.011
- Meyer, M. D., and Pataky, J. K. (2010). Increased severity of foliar diseases of sweet corn infected with maize dwarf mosaic and sugarcane mosaic viruses. *Plant Dis.* 94, 1093–1099. doi: 10.1094/PDIS-94-9-1093
- Mezzalama, M. (2015). MLN Pathogen Diagnosis, MLN-Free Seed Production and Safe Exchange to Non-Endemic Countries CIMMYT.
- 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
- Nault, L. R., Styer, W. E., Coffey, M. E., Gordon, D. T., Negi, L. S., and Niblett, C. L. (1978). Transmission of maize chlorotic mottle virus by chrysomelid beetles. *Phytopathology* 68, 1071–1074. doi: 10.1094/Phyto-68-1071
- Ngugi, H. K., King, S. B., Abayo, G. O., and Reddy, Y. V. R. (2002). Prevalence, incidence, and severity of sorghum diseases in western Kenya. *Plant Dis.* 86, 65–70. doi: 10.1094/PDIS.2002.86.1.65
- Odiyo, O., Njoroge, K., Chemining, G., and Beyene, Y. (2014). Performance and adaptability of doubled haploid maize testcross hybrids under drought stress and non-stress conditions. *Int. Res. J. Agric. Sci. Soil Sci.* 4, 150–158. doi: 10.14303/irjas.2014.055
- Patil, G., Chaudhary, J., Vuong, T. D., Jenkins, B., Qiu, D., Kadam, S., et al. (2017). Development of SNP genotyping assays for seed composition traits in soybean. *Int. J. Plant Genomics* 2017:6572969. doi: 10.1155/2017/6572969
- Patterson, E. L., Fleming, M. B., Kessler, C. K., Nissen, S. J., and Gaines, T. A. (2017). A KASP genotyping method to identify northern watermilfoil, Eurasian watermilfoil, and their interspecific hybrids. *Front. Plant Sci.* 8:752. doi: 10.3389/fpls.2017.00752
- Pemberton, C. E., and Charpentier, L. J. (1969). “Insect Vectors of Sugarcane Virus Diseases,” in *Pests of Sugarcane*. eds. J. R. Williams, J. R. Metcalfe, R. W. Mungomery and R. Mathers (Amsterdam: Elsevier Publishing Co.), 411–425.
- Pereira, G. S., Camargos, R. B., Balestre, M., Pinho, R. G. V., and Melo, W. M. (2015). Indirect selection for resistance to ear rot and leaf diseases in maize lines using biplots. *Genet. Mol. Res.* 14, 11052–11062. doi: 10.4238/2015. September.21.18
- Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., et al. (2016). Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor. Appl. Genet.* 129, 1843–1860. doi: 10.1007/s00122-016-2743-x
- Reddy, M., and Singh, K. (1984). Evaluation of a world collection of chickpea germplasm accessions for resistance to ascochyta blight. *Plant Dis.* 68, 900–901.
- Ribaut, J. M., and Ragot, M. (2007). Marker-assisted selection to improve drought adaptation in maize: the backcross approach, perspectives, limitations, and alternatives. *J. Exp. Bot.* 58, 351–360. doi: 10.1093/jxb/erl214
- SAS institute Inc (2016). SAS® 9.4 Intelligence Platform: Data Administration Guide. 4th Edn. Cary, NC.
- Semagn, K., Babu, R., Hearne, S., and Olsen, M. (2014). Single nucleotide polymorphism genotyping using kompetitive allele specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol. Breed.* 33, 1–14. doi: 10.1007/s11032-013-9917-x
- Sukruth, M., Paratwagh, S. A., Sujay, V., Kumari, V., Gowda, M. V. C., Nadaf, H. L., et al. (2015). Validation of markers linked to late leaf spot and rust resistance, and selection of superior genotypes among diverse recombinant inbred lines and backcross lines in peanut (*Arachis hypogaea* L.). *Euphytica* 204, 343–351. doi: 10.1007/s10681-014-1339-2
- Tanweer, F. A., Rafi, M. Y., Sijam, K., Rahim, H. A., Ahmed, F., Ashkani, S., et al. (2015). Introgression of blast resistance genes (putative pi-b and pi-Kh) into elite rice cultivar MR219 through marker-assisted selection. *Front. Plant Sci.* 6:1002. doi: 10.3389/fpls.2015.01002
- Tivoli, B., Baranger, A., Avila, C. M., Banniza, S., Barbetti, M., Chen, W., et al. (2006). Screening techniques and sources of resistance to foliar diseases caused by major necrotrophic fungi in grain legumes. *Euphytica* 147, 223–253. doi: 10.1007/s10681-006-3131-4

- Vales, M. I., Schön, C. C., Capettini, F., Chen, X. M., Corey, A. E., Mather, D. E., et al. (2005). Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. *Theor. Appl. Genet.* 111, 1260–1270. doi: 10.1007/s00122-005-0043-y
- Xia, Z., Zhao, Z., Chen, L., Li, M., Zhou, T., Deng, C., et al. (2016). Synergistic infection of two viruses MCMV and SCMV increases the accumulations of both MCMV and MCMV-derived siRNAs in maize. *Sci. Rep.* 6:20520. doi: 10.1038/srep20520
- Ziyomo, C., and Bernardo, R. (2013). Drought tolerance in maize: indirect selection through secondary traits versus genomewide selection. *Crop Sci.* 52, 1269–1275. doi: 10.2135/cropsci2012.11.0651

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 Awata, Ifie, Danquah, Jumbo, Suresh, Gowda, Marchelo-Dragga, Olsen, Shorinola, Yao, Boddupalli and Tongoon. 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.



Novel Genomic Regions of *Fusarium* Wilt Resistance in Bottle Gourd [*Lagenaria siceraria* (Mol.) Standl.] Discovered in Genome-Wide Association Study

Yanwei Li, Ying Wang, Xinyi Wu, Jian Wang, Xiaohua Wu, Baogen Wang, Zhongfu Lu and Guojing Li*

Institute of Vegetables, State Key Laboratory for Quality and Safety of Agro-Products, Zhejiang Academy of Agricultural Sciences, Hangzhou, China

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Manjusha Verma,
National Bureau of Plant Genetic
Resources (ICAR), India
Shouvik Das,
Indian Agricultural Research Institute
(ICAR), India
Aamir W. Khan,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

*Correspondence:

Guojing Li
ligj@mail.zaas.ac.cn

Specialty section:

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

Received: 06 January 2021

Accepted: 30 March 2021

Published: 07 May 2021

Citation:

Li Y, Wang Y, Wu X, Wang J,
Wu X, Wang B, Lu Z and Li G (2021)
Novel Genomic Regions of *Fusarium*
Wilt Resistance in Bottle Gourd
[*Lagenaria siceraria* (Mol.) Standl.]
Discovered in Genome-Wide
Association Study.
Front. Plant Sci. 12:650157.
doi: 10.3389/fpls.2021.650157

Fusarium wilt (FW) is a typical soil-borne disease that seriously affects the yield and fruit quality of bottle gourd. Thus, to improve resistance to FW in bottle gourd, the genetic mechanism underlying FW resistance needs to be explored. In this study, we performed a genome-wide association study (GWAS) based on 5,330 single-nucleotide polymorphisms (SNPs) and 89 bottle gourd accessions. The GWAS results revealed a total of 10 SNPs ($P \leq 0.01$, $-\log_{10}P \geq 2.0$) significantly associated with FW resistance that were detected in at least two environments (2019DI, 2020DI, and the average across the 2 years); these SNPs were located on chromosomes 1, 2, 3, 4, 8, and 9. Linkage disequilibrium (LD) block structure analysis predicted three potential candidate genes for FW resistance. Genes *HG_GLEAN_10001030* and *HG_GLEAN_10001042* were within the range of the mean LD block of the marker BGRSe_14202; gene *HG_GLEAN_10011803* was 280 kb upstream of the marker BGRSe_00818. Real-time quantitative PCR (qRT-PCR) analysis showed that *HG_GLEAN_10011803* was significantly up-regulated in FW-infected plants of YD-4, Yin-10, and Hanbi; *HG_GLEAN_10001030* and *HG_GLEAN_10001042* were specifically up-regulated in FW-infected plants of YD-4. Therefore, gene *HG_GLEAN_10011803* is likely the major effect candidate gene for resistance against FW in bottle gourd. This work provides scientific evidence for the exploration of candidate gene and development of functional markers in FW-resistant bottle gourd breeding programs.

Keywords: bottle gourd, *Fusarium* wilt, genome-wide association (GWAS), novel genomic regions, qRT-PCR

INTRODUCTION

Bottle gourd [*Lagenaria siceraria* (Mol.) Standl.] ($2n = 2 \times = 22$), also known as calabash or long melon, is a member of the *Legendaria* genus, Cucurbitaceae family, and is an annual plant (Whitaker, 1971; Erickson et al., 2005). Its fresh young fruits are eaten as a vegetable in parts of Asia and Africa, while dry old fruits are used for containers, musical instruments, and crafts (Heiser, 1979; Morimoto and Mvere, 2004). Moreover, due to its strong resistance to disease and abiotic

stress, bottle gourd is commonly used as grafting rootstock for other crops, such as watermelon and melon (Lee, 1994; Yetisir and Sari, 2003). According to records, bottle gourd has been cultivated in China for more than 7,000 years, covering a cultivation area of more than 2 million mu; this is an important vegetable crop in China, especially in the southern part.

Fusarium wilt (FW), which is caused by *Fusarium oxysporum*, is a typical soil-borne disease of economic crops worldwide (Katan, 1994; Wechter et al., 2012; Bodah, 2017). Since this pathogen can survive in the absence of host-infected plants, once the disease occurs in the field, *F. oxysporum* is likely to remain in the soil indefinitely, which seriously affects the yield of crops (Cha et al., 2016; Khan et al., 2017). FW has a broad host on cucurbit crops, including watermelon, melon, cucumber, luffa, and bottle gourd (Bodah, 2017). Usually, the infected plants morphologically show a constriction in the stem xylem, resulting in vascular bundle clogging, plant wilting, or death (Freeman et al., 2002; Singh et al., 2017). FW commonly occurs during the whole growth period of bottle gourd, especially the flowering to fruiting period. Its incidence is approximately 20–40%, and severe cases could cause devastating losses (data from Ningbo Institute of Agriculture), which severely restrict the sustainable development of the bottle gourd industry.

Breeding resistant varieties is one of the most effective and economic methods to control FW disease. At present, a series of commercial varieties that are highly resistant to FW has been grown for production, such as watermelon, melon, and cowpea (Zink and Gubler, 1985; Martyn and Bruton, 1989; Ehlers et al., 2000, 2009). To improve FW resistance, we need to exploit markers tightly linked to FW resistance using quantitative trait loci (QTL) and then generate germplasm by molecular marker assistant selection (MAS; Zhao et al., 2014; Li et al., 2017). To date, QTLs/genes conferring FW resistance have been thoroughly studied in many crops. For example, a dominant gene *I-2* that confers resistance to race 2 of FW was cloned in tomato by map-based cloning (Simons et al., 1998; Catanzariti et al., 2015). Using the same map-based cloning technique, Joobeur et al. (2004) identified two candidate genes of melon FW resistance in the physical range of Fom-2. Several genes associated with cowpea FW resistance were identified using QTL analysis in “California Blackeye 27” (Pottorff et al., 2012, 2013). In cucumber, a major QTL Foc2.1 conferring resistance to FW was detected using recombinant inbred lines (Zhang et al., 2014), and a major QTL, Fo-1.1, associated with FW resistance to race 1 was identified by using selective genotyping in genetic populations derived from related watermelon lines (Lambel et al., 2014). However, research progress on the FW resistance of bottle gourd is relatively limited. Only its specialization of *F. oxysporum* f. sp. *lagenariae* has been reported, whereas the genetic mechanism of FW resistance and related genes/QTLs are unknown in bottle gourd.

To date, a high-density genetic map has been constructed, and a series of ISSR, SSR, and single-nucleotide polymorphism (SNP) markers has been exploited for bottle gourd (Xu et al., 2011, 2014; Bhawna et al., 2014), allowing the establishment of various marker–trait associations, such as association analysis for the free glutamate content of bottle gourd (Wu et al., 2017). Genome-wide association study (GWAS), based on linkage disequilibrium

(LD), has also been widely used in the study of plants, and various results have been reported (Joobeur et al., 2004; Wang et al., 2009; Sabbavarapu et al., 2013; Zhang et al., 2014). In bottle gourd molecular breeding, Wu et al. (2017) performed a GWAS for SNPs related to the free glutamate content of the umami factor. With the development of quantitative genetics, many researchers have proposed different analytical models, such as efficient mixed-model association (Kang et al., 2008), compressed mixed linear model (Zhang et al., 2010), restricted two-stage multi-locus GWAS (He et al., 2017), etc. Among them, general linear model (GLM) and mixed linear model (MLM) are still the common GWAS methods in plants (Huang et al., 2010; Li et al., 2013; Fang et al., 2017). In this study, we initially genotyped 89 bottle gourd accessions using 5,330 SNPs and surveyed the disease index (DI) of FW resistance in two consecutive years. We then performed a GWAS to identify significant associated SNPs and potential candidate genes. Finally, three candidate genes associated with FW resistance were verified by quantitative real-time PCR (qRT-PCR). Our study is the first to use GWAS to identify genomic regions and candidate genes associated with FW resistance. The GWAS results can lay a foundation for MAS breeding and the genetic mechanisms of FW resistance in cucurbit crops.

MATERIALS AND METHODS

Plant Materials

Germplasm consisting of 89 bottle gourd accessions was collected, consisting of 87 accessions from wide areas across China, one accession from Europe, and one accession from Mexico (**Supplementary Table 1**). All accessions (inbred lines) were evaluated for FW resistance in a glasshouse of the Haining Experimental Station (30° N, 120° E) in 2019 and 2020. According to a completely randomized block design, the plants were studied based on two replications in both years.

Inoculation System of FW Resistance in Bottle Gourd

During 2018–2019, bottle gourd FW fungus was isolated from wilted plants that were collected from severely affected areas such as Shaoxing and Haining (**Supplementary Figures 1A,B**). According to the conventional tissue separation method, FW strains with obvious antagonistic effects were obtained by using potato dextrose agar (containing 100 mg/ml Kana and Amp) for screening four to five times (**Supplementary Figure 1C**). Under a microscope with 10 × 40 magnification, small conidia were observed to be ovoid or kidney-shaped, and large conidia were spindle-shaped or sickle-shaped with hooked apex (**Supplementary Figure 1D**). PCR assays showed that the similarity between the sequence of FW isolates and the 16S rRNA sequence was as high as 99%. After cytological tests and PCR detection, the isolates were identified as *F. oxysporum* f. sp. strains and were stored at 4°C at the Zhejiang Academy of Agricultural Sciences, Hangzhou, China.

Each FW strain was shake-cultured on potato sucrose broth for 3 days in the dark at 28°C at 200 rpm. With the use of

a hemacytometer, the conidial suspension was adjusted to a final concentration of 1.0×10^6 conidia/ml with sterile distilled H_2O . The seeds of each accession were sown in mixed soil (nutritional soil/vermiculite = 3: 1) in plastic pots (6 by 6 by 5 cm) and were grown in a glasshouse set at 24°C, 16-h light/18°C, 8-h darkness, 60% humidity. At the second true leaf of the seedling spreading stage, we used the root dipping method for bottle gourd FW resistance screening and testing. Each accession consisted of 10–12 seedlings, and two duplicates were set per environment.

Disease Assessment and Statistical Analysis

Leaf damage was used as a main index to evaluate resistant/susceptible phenotypic traits. The standard reported by Gao et al. (2015) and Xu et al. (2016) was further improved and implemented with a few modifications. We classified the phenotypes of plants according to a 0–4 scale as follows: level 0, no disease symptoms, i.e., immune (I); level 1, slight disease symptoms, featured by less than 25% of leaves with disease symptoms, with normal plant growth, i.e., highly resistant (HR); level 2, slight wilt symptoms, featured by 25–50% of leaves with disease symptoms, i.e., resistant (R); level 3, moderate wilt symptoms, featured by 50–90% of leaves with disease symptoms, i.e., susceptible (S); and level 4, completely wilted or dead plants, i.e., highly susceptible (HS; **Supplementary Figure 2**). After 10–12 replicates per material were evaluated individually, we calculated the mean value to determine the disease severity for each accession. The DI was calculated according to the following equation:

$$DI = \frac{\sum P_i \times n_i}{N \times P_4} \times 100\%,$$

where DI is the disease index, P_i is the grade of the DI, n_i is the plant number of the corresponding DI grade, N is the total number of plants investigated, and P_4 is the highest DI grade.

According to the DI scores, the FW resistance of each material was determined following Shen and Li (2008) with a few modifications: immune ($DI = 0$, level 0), highly resistant ($0 < DI \leq 15\%$, level 1), resistant ($15\% < DI \leq 45\%$, level 2), susceptible ($45\% < DI \leq 75\%$, level 3), and highly susceptible ($DI > 75\%$, level 4).

SNP Genotyping, LD, and Population Structure

The SNP markers used in this study were generated from RAD sequencing with paired-end sequencing (150 bp) on the Illumina HiSeq platform. We initially found 89 bottle gourd accessions that aligned to the Hangzhou gourd reference genome of ~330 Mb (Wang et al., 2018)¹ and then removed those SNPs with a minor allele frequency (MAF) of ≤ 0.01 and a heterozygous rate $\geq 25\%$ for data filtering. This left a total of 6,222 high-quality SNPs. Of these SNPs, 85.66% were located on the 11 chromosomes of the bottle gourd, leaving 5,330 high-quality SNPs. These were used for the correlation analysis of traits

(Wu et al., 2017). The density of SNP markers was estimated to be one SNP per 59.37 kb for the 11 bottle gourd chromosomes.

The LD parameters (r^2) for estimating the LD distance of the genome between pairwise SNPs (MAF > 0.01) were calculated using PLINK V1.07 (Purcell et al., 2007; Xu et al., 2021, unpublished), and the average LD decay was drawn with R. The population structure was constructed by STRUCTURE 2.3.4 software (Evanno et al., 2005). K (number of subgroups) values were set to 2–8, with 10,000 (MOMC, Markov chain Monte Carlo)–100,000 runs (MCMC, Monte Carlo Markov Chain) with four replications. Then, the best value of K was determined by $Ln P(D)$ and ΔK according to the principle of maximum likelihood (Evanno et al., 2005). The neighbor-joining tree was constructed using PHYLIP software. The kinship matrix was assessed based on the SNP dataset using TASSEL 5.2.14 to determine the relatedness among individuals (Anderson and Weir, 2007; Zhang et al., 2010). In previous studies, the population was divided into two subgroups depending on the markers used in the tests (Wu et al., 2017).

Genome-Wide Association Analysis and LD Block Construction

For natural populations, the population structure and relative kinship always lead to high levels of false positives in association maps (Yu et al., 2006). After assessment of the population structure (Q), relative kinship (K), and principal component analysis (PCA) of 89 accessions, four correlation analysis models including (1) a general linear model GLM (Q), GLM (PCA) and (2) a mixed linear model MLM ($Q + K$), MLM (PCA + K) were used to conduct a genome-wide correlation analysis of FW resistance using TASSEL 5.2.14 (Anderson and Weir, 2007; Zhang et al., 2010). The significance threshold for SNP–trait associations was determined by $1/n$, where n is the number of markers in the association panel (Yang et al., 2014), and $P \leq 1/5,330$ or $-\log_{10}P \geq 3.7$. Considering that population structure and kinship reduced the detection efficiency of SNPs associated with FW resistance, the $-\log_{10}P$ value of significantly associated SNPs identified in this study was low, which has also appeared in previous studies (Atwell et al., 2010; Huang et al., 2010). In order to fully exploit the valuable genetic information in the bottle gourd germplasm population, the significant threshold for SNP–trait associations was set as $-\log_{10}P = 2$. This threshold has already been applied to other traits in an association analysis (Li et al., 2015; Zhang et al., 2018). The correlation analysis results were plotted using a Manhattan plot and Q–Q plot based on the “qqman” package in R software.

The genome-wide LD decay rate, defined as the LD block distance where the LD decays to half of its maximum value, was about 400 kb in a natural population of bottle gourd (from Xu et al., 2021, unpublished). We defined the average LD decay distance as the candidate region to select candidate genes associated with large-effect SNPs. The genome of “Hangzhou gourd” was used as a reference sequence (Wang et al., 2018). Based on the genomic annotations of GourdBase,² potential candidate genes for FW resistance were predicted.

¹<http://www.gourdbase.cn>

²<http://www.gourdbase.cn>

Validation of Candidate Genes

The expression levels of the candidate genes were measured before and after infecting plants with FW by using qRT-PCR. Based on the phenotype data in 2019 and 2020, Hanbi (HR to FW, level 1), Yin-10 (HR to FW, level 1), and YD-4 (HS to FW, level 4) were chosen as extreme materials and were cultivated in the glasshouse. The leaves from healthy plants (CK) and treatment plants were collected 3 days after FW infection and stored in liquid nitrogen. Total RNA was extracted from Hanbi, Yin-10, and YD-4 leaves using an RNA Simple Total RNA kit (Tiangen, China). After the quality and concentration of total RNA were evaluated using 1% agarose gel and an Agilent 2100 Bioanalyzer, complementary DNA (cDNA) was synthesized by using a Script cDNA Kit (Tiangen, China) with a standard protocol. The CDS sequences of genes were obtained from the GourdBase website.³ qRT-PCR primers (Supplementary Table 2) were designed using the Integrated DAN Technologies website⁴ and were synthesized by Sangon Biotech (Shanghai) Co., Ltd. The bottle gourd *TuB-α* gene (*BG_GLEAN_10019523*) was used as the internal control gene. qRT-PCR was performed on a Bio-Rad CFX96 Touch q-PCR System (Bio-Rad, CA, United States) with SuperReal PreMix Plus/SYBR Green (Tiangen, China). Each reaction was replicated three times. The relative expression level of candidate genes was evaluated by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001); healthy plants (CK) served as the control.

³<http://www.gourdbase.cn>

⁴<https://www.idtdna.com/scitools/Applications/RealTimePCR/Default.aspx>

Student's *t*-test was used for statistical analyses (* $0.01 \leq P < 0.05$, ** $0.001 \leq P < 0.01$, *** $P < 0.001$).

RESULTS

Identification of a *F. oxysporum* f. sp. *lagenariae* Race

According to the conventional tissue separation method, purified strains from *Fusarium* wilt-infected plants were obtained. Through morphological identification of the colony, the microscopic view of its conidia, and PCR detection of its sequence (Supplementary Figure 1), we preliminarily identified the bottle gourd wilt isolates as *F. oxysporum* f. sp. Due to differences in the infectivity and pathogenicity of different strains to cucurbit crops, individual strains of *F. oxysporum* usually infect only one or few host species. Thus, to better distinguish the different races of *F. oxysporum* f. sp., we still relied on the special host for identification. The pathogenicity results showed that bottle gourd plants had obvious wilt infection symptoms, featured by the first and second leaves that were more than 50% wilted and the third and fourth leaves that were crumpled. However, there were no symptoms of wilt infection in watermelon, melon, cucumber, and luffa plants (Figure 1). Therefore, we proposed that the isolated *F. oxysporum* f. sp. was a *F. oxysporum* f. sp. *lagenariae* race and was named physiological race ShaoX-1, which was used for the subsequent phenotypic identification of bottle gourd.

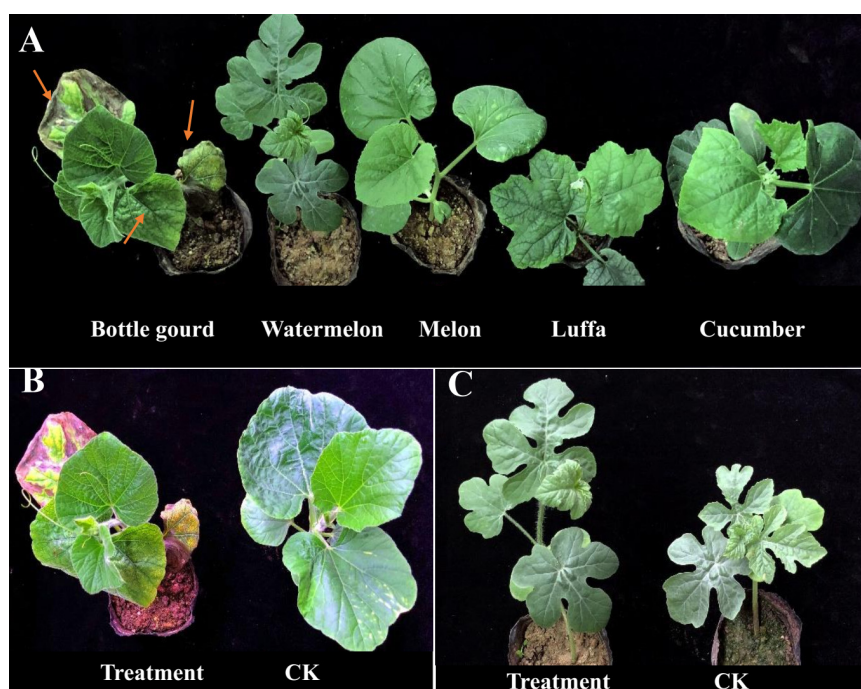


FIGURE 1 | Morphology and *Fusarium* wilt (FW)-inoculated infestation response of cucurbit crops. **(A)** Bottle gourd was more susceptible than other crops after FW inoculation. **(B)** Plant phenotype of bottle gourd inoculated treatment group and control group after 10 days. **(C)** Plant phenotype of watermelon inoculated treatment group and control group after 10 days.

Phenotypic Analysis of FW Resistance in the Natural Population

In this study, a total of 89 bottle gourd accessions were evaluated for resistance to *Fusarium* wilt in a glasshouse in 2019 (DI2019) and 2020 (DI2020), with two replicates per environment. DI, as an evaluation of FW resistance, had a wide range of phenotypic variation in the 2-year trials. The DI of all accessions ranged from 6 to 95%, with a mean value of 46% in 2019 (DI2019), and from 11 to 94%, with a mean value of 55% in 2020 (DI2020). The ANOVA results showed that the broad-sense heritability (h^2) was 87.19% across the 2 years (Table 1), suggesting that the genetic effect played a predominant role in the phenotypic variation of FW resistance in bottle gourd. We divided the DI into five levels (Supplementary Figure 2): immune (level 0), highly resistant (level 1), resistant (level 2), susceptible (level 3), and highly susceptible (level 4), according to relevant previous studies (Gao et al., 2015; Xu et al., 2016). Only a tiny percentage of accessions had DI values less than 15% (8 in 2019 and 1 in 2020), whereas the majority of the accessions were within the range of 15.01–45% (35 in 2019 and 28 in 2020) and 45.01–75% (37 in 2019 and 42 in 2020). When DI exceeded 75%, there were 9 accessions in 2019 and 18 accessions in 2020 (Figure 2). Unfortunately, we did not select for any material that was immune to FW in the 2-year trials; only a small amount of material had high resistance to FW (Figure 2), which showed that the bottle gourd germplasm resource of FW resistance is scarce. The correlation coefficient between the 2-year trials was as high as 0.62 (Supplementary Figure 3), and the frequency distribution of DI was approximately normally distributed, which indicated that this natural population could be suitable for correlation analysis for FW resistance.

SNP Marker Analysis

The SNP markers used in this work resulted from RAD-sequencing by using the Illumina HiSeq platform. After removal of SNPs with a MAF of ≤ 0.01 and a heterozygous rate $\geq 25\%$, a total of 5,330 high-quality SNPs were retained for GWAS of the FW resistance trait. These SNPs covered all 11 chromosomes, with an uneven distribution across the genome (Table 2). The average density of SNP markers was about 59.37 kb/SNP. The maximum marker density was found on chromosome 11 (101.18 kb/SNP) followed by chromosome 6 (67.25 kb/SNP), whereas the minimum marker density was found on chromosome 1 (42.11 kb/SNP). Based on the SNP markers, we estimated the population structure of 89 bottle gourds using STRUCTURE software and cluster analysis. The

TABLE 1 | Descriptive statistics and heritability (h^2) of the *Fusarium* wilt disease index.

| Trial | Maximum | Minimum | Mean | SD | CV (%) | Heritability (%) |
|--------|---------|---------|------|------|--------|------------------|
| DI2019 | 0.95 | 0.06 | 0.46 | 0.22 | 48.77 | 87.19 |
| DI2020 | 0.94 | 0.11 | 0.55 | 0.22 | 41.21 | |
| Mean | 0.94 | 0.11 | 0.50 | 0.21 | 42.29 | |

SD, standard deviation; CV (%), coefficient of variation.

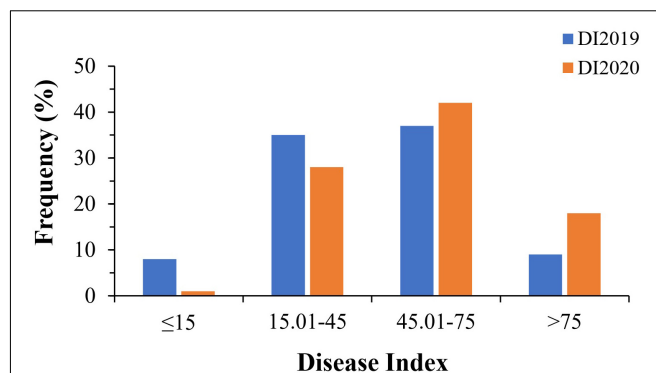


FIGURE 2 | Phenotypic distribution of *Fusarium* wilt resistance in 89 bottle gourd accessions.

TABLE 2 | Single-nucleotide polymorphism (SNP) marker distribution on 11 chromosomes of bottle gourd.

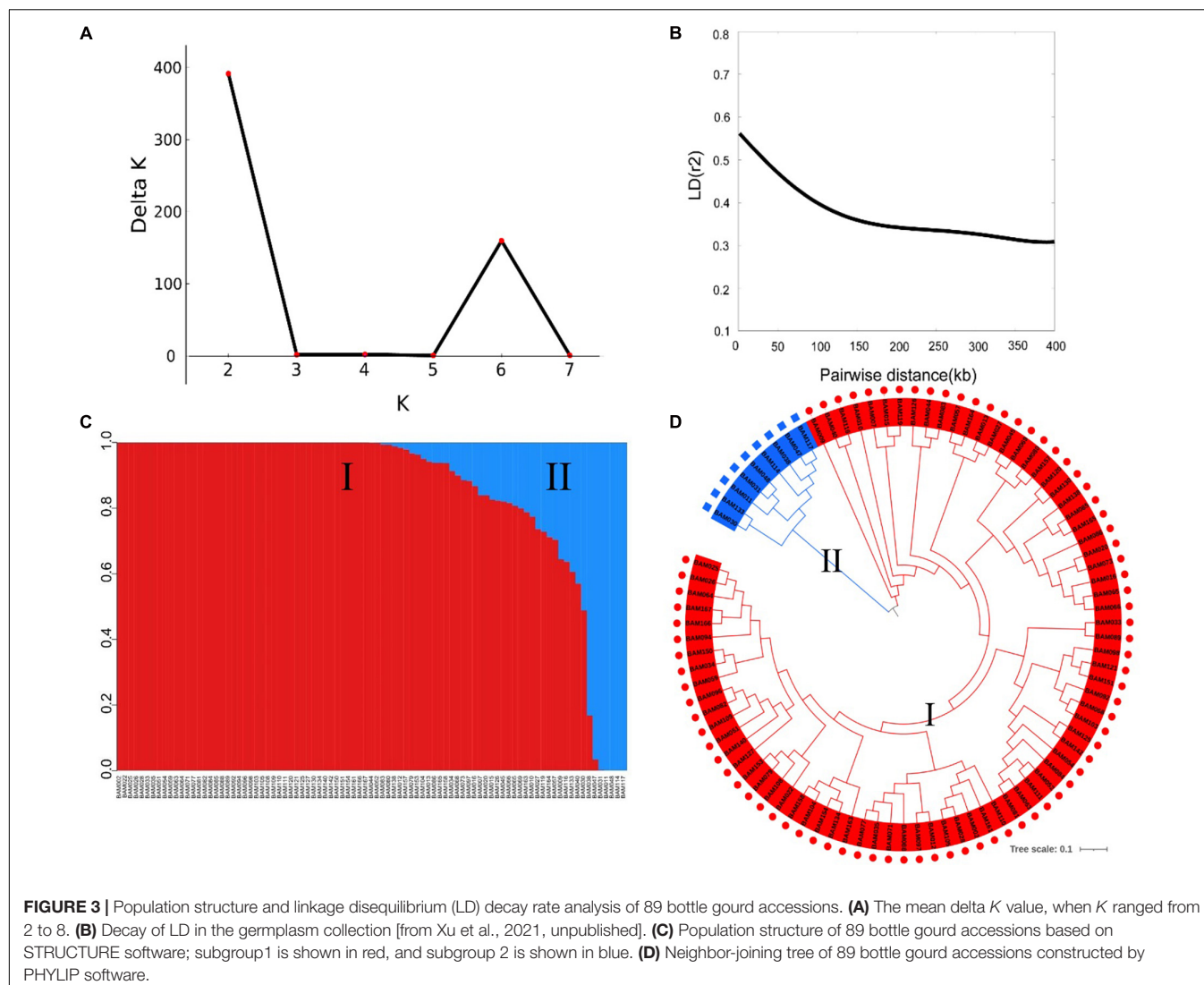
| Chr. | Chromosome length (Mb) | Number of SNP | Density of SNP (kb/SNP) | PIC |
|-------|------------------------|---------------|-------------------------|------|
| chr1 | 28.39 | 674 | 42.11 | 0.10 |
| chr2 | 29.76 | 631 | 47.17 | 0.14 |
| chr3 | 30.34 | 563 | 53.90 | 0.13 |
| chr4 | 32.30 | 637 | 50.70 | 0.11 |
| chr5 | 35.14 | 553 | 63.55 | 0.12 |
| chr6 | 26.83 | 399 | 67.25 | 0.12 |
| chr7 | 23.92 | 389 | 61.49 | 0.13 |
| chr8 | 23.22 | 505 | 45.98 | 0.10 |
| chr9 | 19.99 | 370 | 54.03 | 0.12 |
| chr10 | 26.30 | 400 | 65.75 | 0.10 |
| chr11 | 21.15 | 209 | 101.18 | 0.12 |

delta K reached a sharp peak when K was 2. Therefore, this association population was divided into two subgroups, namely, subgroup 1 and subgroup 2 (Figures 3A,C). Subgroup 1 contained 80 accessions, and subgroup 2 was small and included nine accessions. A neighbor-joining result also classified the population into two subgroups, consistent with the population grouping result (Figure 3D). Because all accessions have some distant relationship, there were no primary factors, such as geographic distribution, affecting the population structure of the 89 accessions. Genotype data were subjected to correlation analysis of the free glutamate content trait in bottle gourd (Wu et al., 2017).

To determine the mapping resolution of the GWAS, we estimated the genome-wide LD decay distance of the association population. The average LD decay distance was approximately 400 kb, where $r^2 = 0.3$ for all chromosomes [half of its maximum value, from Xu et al., 2021, unpublished] (Figure 3B).

Model Comparison for Correlation Analysis

To reduce a false positive association, we applied four kinds of association analysis models to GWAS for FW resistance in bottle



gourd. Based on the mean value of DI across the 2 years, quantile-quantile (Q-Q) plots were drawn (**Supplementary Figure 4**). The results showed that there was a large deviation in the $-\log_{10}P$ value between the observed values and the expected values in GLM (PCA) and GLM (Q) models, which indicated that the two models might cause a high false positive correlation. Due to the introduction of the covariable K , the observed $-\log_{10}P$ values fit well with the expected values in the MLM (PCA + K) and MLM (Q + K) models, indicating that those two models could effectively control the false positive of the association analysis results. Taking into account the Q-Q plots of each environment, the MLM (Q + K) model (red scatter plot in **Supplementary Figure 4**) was selected for the follow-up association analysis for FW resistance.

Genome-Wide Association Analysis

A GWAS was performed to detect SNPs associated with FW resistance between 5,330 SNP markers and 89 phenotype data points from the mean across the 2 years (aDI) and within an individual year (DI2019 and DI2020). The Manhattan plots

and Q-Q plots for the GWAS results are shown in **Figure 4**. The GWAS result showed that 20 SNPs (with a significance threshold of $p \leq 0.01$, $-\log_{10}P \geq 2.0$) significantly associated with FW resistance were detected in at least one environment (**Supplementary Table 3**), including 12 SNPs from the 2019 data, 11 SNPs from the 2020 data, and 11 SNPs from the mean data. Among these SNPs, 10 significantly correlated SNP sites were detected in at least two environments, which were located on chromosomes 1, 2, 3, 4, 8, and 9, indicating that the FW resistance of bottle gourd is controlled by multiple genes (**Figures 4A–C**). The phenotypic variation explained by these sites ranged from 8.82 to 15.03% (**Table 3**). Among them, markers of BGRSe_14212 and BGRSe_14202 were located on chromosome 9, and those two SNP markers were within the range of the genome-wide LD block (400 kb). BGRSe_14202 was detected in all three environments with relatively high significant levels ($-\log_{10}P = 2.81/2.49/2.46$) and an effect on FW ($R^2 = 14.14\%/13.90\%/10.49\%$). Therefore, the region range of chromosome 9 may contain the major genes associated with FW resistance. On chromosome 8, two

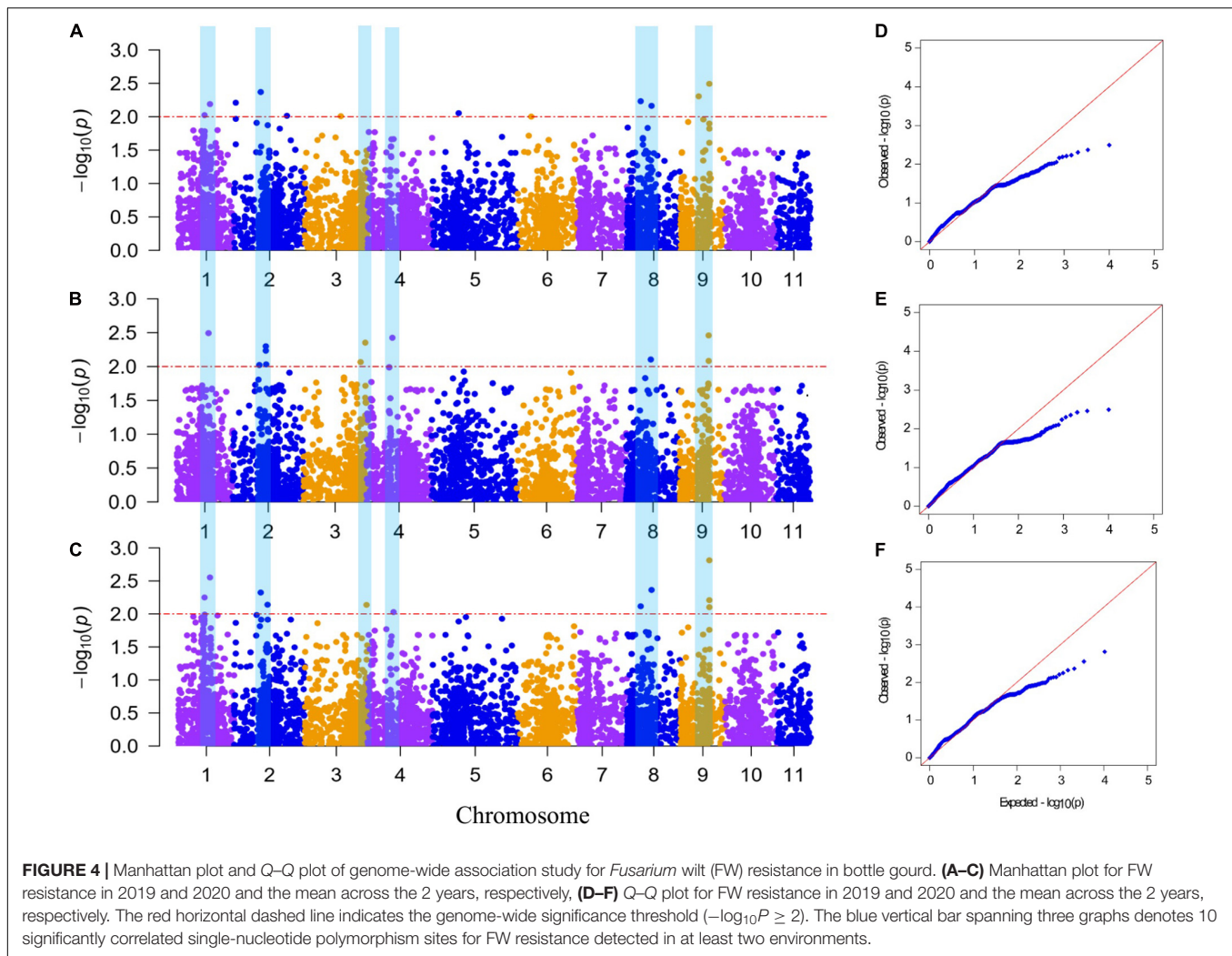


FIGURE 4 | Manhattan plot and Q-Q plot of genome-wide association study for *Fusarium* wilt (FW) resistance in bottle gourd. **(A–C)** Manhattan plot for FW resistance in 2019 and 2020 and the mean across the 2 years, respectively, **(D–F)** Q-Q plot for FW resistance in 2019 and 2020 and the mean across the 2 years, respectively. The red horizontal dashed line indicates the genome-wide significance threshold ($-\log_{10}P \geq 2$). The blue vertical bar spanning three graphs denotes 10 significantly correlated single-nucleotide polymorphism sites for FW resistance detected in at least two environments.

SNP markers were detected with a certain LD distance away. BGRSeSe_12911 was significantly correlated with FW resistance in all three environments, and BGRSeSe_12338 was detected in DI2019 and aDI. Two SNP markers, BGRSeSe_02569 and BGRSeSe_02108, were detected on chromosome 2. Of these two, BGRSeSe_02569 explained the largest phenotypic variation in DI2020 and aDI, *i.e.*, 16.19 and 15.38%, respectively. BGRSeSe_02108 was detected in three environments, with a contribution rate for phenotypic variation of 12.60, 11.03, and 11.28%. BGRSeSe_01042 and BGRSeSe_00818 were located on chromosome 1. One of the markers, BGRSeSe_00818 ($-\log_{10}P = 2.25/2.02$), was significantly correlated with FW resistance in the two environments of DI2019 and aDI, and its contribution rate for phenotypic variation was 12.26 and 12.84%, respectively (Table 3).

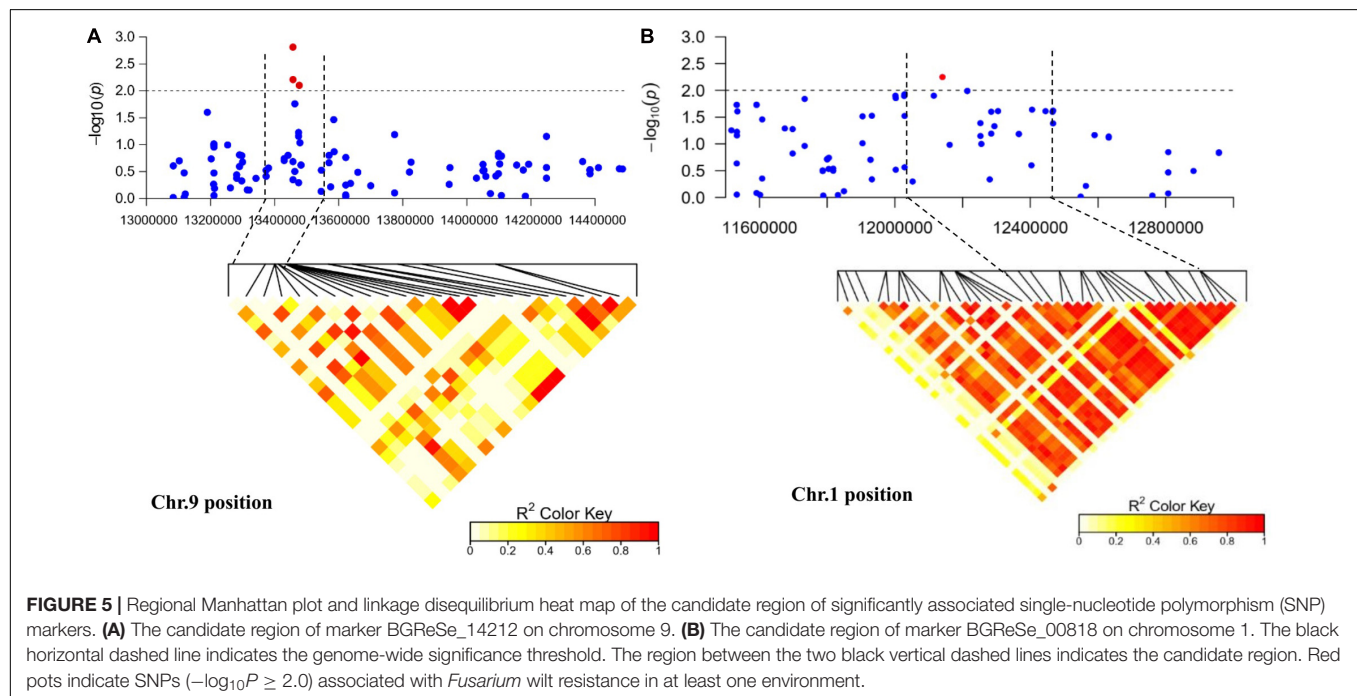
Prediction of Candidate Genes for FW Resistance

In this study, we were interested in the markers with the greatest effect, such as marker BGRSeSe_00818 (MAF = 1.16) on chromosome 1 and markers BGRSeSe_14202 (MAF = 1.09) and BGRSeSe_14212 (MAF = 1.07) on chromosome 9. To

reduce the range of candidate regions, we performed LD block structure analysis. The results showed that BGRSeSe_14202 and BGRSeSe_14212 could form an obvious LD (± 400 kb) block, meaning that these two SNPs were closely linked (Figure 5A). The candidate region of chromosome 9 was narrowed down to approximately 140 kb. This region contained 15 genes, of which two candidate genes were significantly associated with FW resistance of bottle gourd (Table 4). Both of them, *HG_GLEAN_10001030* (ethylene-responsive transcription factor RAP2) and *HG_GLEAN_10001042* (GDSL esterase), are involved in signaling pathways, such as resistance genes and hormone induction. LD block reduced the candidate region of BGRSeSe_00818 to about 415 kb, which contained seven genes (Figure 5B). Among them, *HG_GLEAN_10011803* encodes carboxylesterase and a CDPK-related kinase protein and plays a role in the signal transduction pathway. To confirm whether the potential candidate genes participated in the FW resistance pathway, the expression patterns of the three genes in both FW-infected and healthy bottle gourd plants were analyzed *via* qRT-PCR. The representative materials were selected from the association analysis population in this study. The DI of Yin-10 and Hanbi was 10.83 and 13.44%, respectively, and both were

TABLE 3 | Significant markers associated with *Fusarium* wilt resistance in at least two environments.

| Marker | Chromosome | Positive | Allelic | 2019DI | | 2020DI | | aDI | |
|--------------|------------|------------|---------|---------------|-----------|---------------|-----------|---------------|-----------|
| | | | | $-\log_{10}P$ | R^2 (%) | $-\log_{10}P$ | R^2 (%) | $-\log_{10}P$ | R^2 (%) |
| BGReSe_14212 | 9 | 13,476,930 | A/G | | | 2.10 | 9.76 | 2.08 | 8.82 |
| BGReSe_14202 | 9 | 13,457,203 | A/G | 2.81 | 14.14 | 2.49 | 13.90 | 2.46 | 10.49 |
| BGReSe_12911 | 8 | 11,449,774 | A/G | 2.36 | 10.14 | 2.16 | 10.31 | 2.10 | 10.33 |
| BGReSe_12338 | 8 | 6,378,304 | C/T | 2.23 | 15.03 | | | 2.12 | 12.87 |
| BGReSe_5941 | 4 | 12,071,538 | C/T | | | 2.42 | 13.60 | 2.03 | 9.88 |
| BGReSe_5382 | 3 | 28,668,323 | A/G | | | 2.35 | 15.40 | 2.14 | 12.25 |
| BGReSe_2569 | 2 | 15,601,788 | A/G | | | 2.14 | 16.19 | 2.03 | 15.38 |
| BGReSe_2108 | 2 | 12,417,989 | A/G | 2.37 | 12.60 | 2.32 | 11.03 | 2.02 | 11.28 |
| BGReSe_1042 | 1 | 14,684,871 | C/T | 2.55 | 11.06 | 2.49 | 12.83 | 2.19 | 12.30 |
| BGReSe_818 | 1 | 12,140,445 | C/T | 2.25 | 12.26 | | | 2.02 | 12.84 |



highly resistant (HR, level 1) to FW. The DI of YD-4 was 87.81%, i.e., highly susceptible (HS, level 4) to FW. The expression pattern of three potential candidate genes *HG_10011803*, *HG_10001030*, and *HG_10001042* in materials YD-4, Yin-10, and Hanbi is presented (Figure 6). Compared to healthy YD-4 (HS material, level 4), the expression levels of the three candidate genes were all significantly higher ($P < 0.001$) in the FW-infected group (3 days after infection) (Figure 6A). For Yin-10 and Hanbi (HR materials, level 1), the expression level of gene *HG_10011803* showed a significant difference ($P < 0.05$ and $P < 0.001$) between FW-infected and healthy groups. However, the expression levels of *HG_10001030* and *HG_10001042* in infected plants showed a higher or lower expression level than those in healthy Yin-10/Hanbi plants, without a significant difference (Figures 6B,C). Combining the above-mentioned interesting results, we speculated that *HG_10011803* is a major effect gene, whereas *HG_10001030* and *HG_10001042* might be

the candidate genes involved in the FW resistance response in bottle gourd.

DISCUSSION

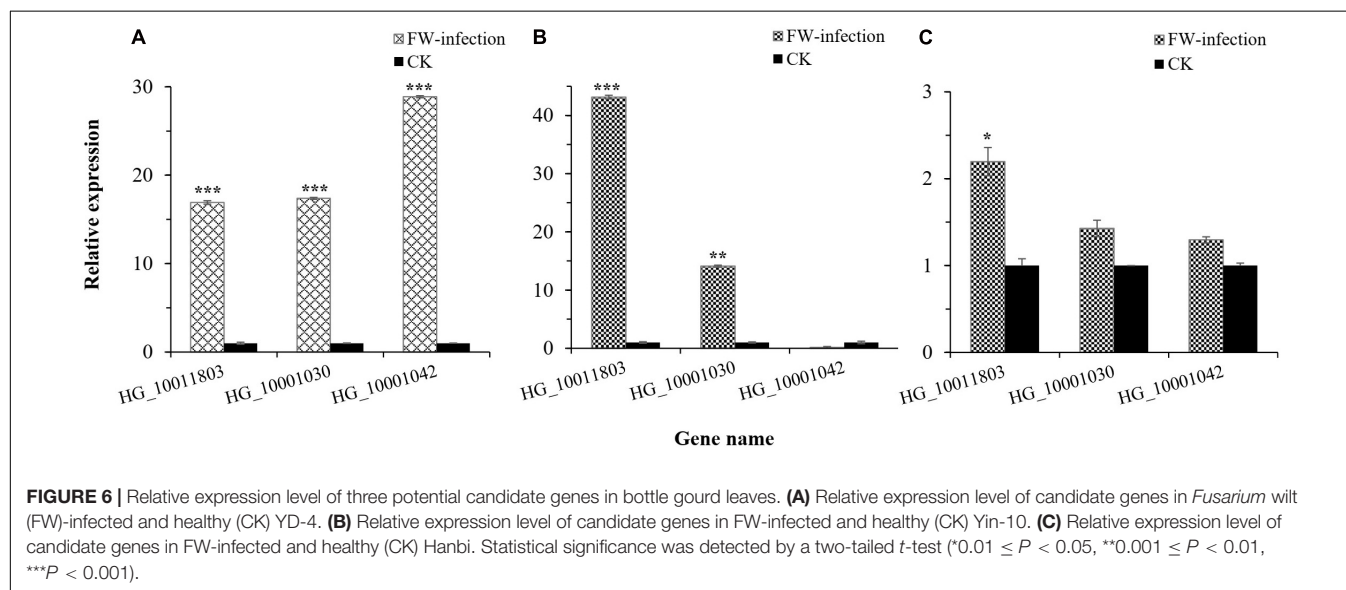
Fusarium wilt is one of the most important diseases throughout the world, which seriously affects the yield and quality of cucurbit crops (Miguel et al., 2004; Wechter et al., 2012; Oumouloud et al., 2013). The genetic mechanism of resistance to FW in cucurbit crops is complex, showing genetic diversity. However, there are no studies on the genetic effect and inheritance of genes governing FW resistance in bottle gourd, and molecular markers linked to FW resistance are also poorly reported.

Genome-wide association study has emerged as a powerful tool to study complex traits and genetic variations in SNP loci and has been successfully applied to different crops in recent

TABLE 4 | Function annotation and genes in candidate intervals of *Fusarium* wilt resistance single-nucleotide polymorphisms.

| Gene model | Chromosome | Start (bp) | End (bp) | Gene Ontology biological process descriptions |
|--------------|------------|------------|------------|--|
| HG_10001024 | 9 | 13,370,905 | 13,396,585 | Chloroplastic/mitochondrial isoform X1 |
| HG_10001026 | 9 | 13,403,689 | 13,406,203 | NA |
| HG_10001028 | 9 | 13,411,513 | 13,414,862 | Mitochondrial dicarboxylate/tricarboxylate transporter |
| HG_10001029 | 9 | 13,422,565 | 13,425,196 | Tubulin alpha-3 chain |
| HG_10001030* | 9 | 13,426,782 | 13,427,351 | Ethylene-responsive transcription factor |
| HG_10001031 | 9 | 13,435,260 | 13,440,011 | Importin |
| HG_10001032 | 9 | 13,440,450 | 13,443,769 | Importin-5 |
| HG_10001034 | 9 | 13,447,396 | 13,453,363 | Chloroplastic isoform X1 |
| HG_10001035 | 9 | 13,453,903 | 13,460,368 | Indole-3-acetaldehyde oxidase-like isoform X1 |
| HG_10001037 | 9 | 13,467,477 | 13,475,348 | Indole-3-acetaldehyde oxidase-like |
| HG_10001040 | 9 | 13,491,070 | 13,493,857 | NADH-cytochrome b5 reductase 1-like isoform X3 |
| HG_10001041 | 9 | 13,494,064 | 13,495,437 | Chloroplastic-like isoform X1 |
| HG_10001042* | 9 | 13,495,614 | 13,500,201 | GDSL esterase |
| HG_10001043 | 9 | 13,505,538 | 13,506,317 | Protein YLS9-like |
| HG_10001044 | 9 | 13,510,514 | 13,520,956 | La-related protein 1A |
| HG_10011790 | 1 | 12,064,283 | 12,065,568 | Probable carboxylesterase 15 |
| HG_10011791 | 1 | 12,073,500 | 12,084,249 | Protein TRANSPARENT TESTA 12-like |
| HG_10011796 | 1 | 12,254,943 | 12,255,467 | Hypothetical protein |
| HG_10011797 | 1 | 12,274,175 | 12,283,744 | Hypothetical protein |
| HG_10011798 | 1 | 12,346,824 | 12,347,511 | Hypothetical protein |
| HG_10011799 | 1 | 12,379,221 | 12,379,829 | Zinc-finger homeodomain protein-like |
| HG_10011803* | 1 | 12,423,049 | 12,429,778 | CDPK-related kinase 1-like isoform X1 |

*potential candidate genes for FW resistance.



years (Joobeur et al., 2004; Wang et al., 2009; Sabbavarapu et al., 2013; Zhang et al., 2014). Especially the general linear model and mixed linear model are still the common GWAS methods in plants (Huang et al., 2010; Li et al., 2013; Fang et al., 2017). GLM, based on a linear regression model, is usually used for the analysis of quantitative traits and discrete resistance traits. MLM, based on population structure (*Q*) and kinship (*K*) as covariance, could better reduce the false positive association (Yu et al., 2006). Taking into account the deviation between

expected $-\log_{10}P$ and observed $-\log_{10}P$ in Q-Q plots, we finally selected the MLM (*Q* + *K*) (Supplementary Figure 4) as GWAS model for FW resistance. In this study, we used a GWAS to evaluate a population of 89 accessions for FW resistance under glasshouse inoculation conditions. A total of 20 SNPs ($P \leq 0.01$, $-\log_{10}P \geq 2.0$) significantly associated with FW resistance were identified in at least one environment (Supplementary Table 3). These sites were distributed on seven chromosomes, which could explain the phenotypic variation

up to 16.19%. Among them, 10 significantly correlated SNP sites were detected in at least two environments, which were located on chromosomes 1, 2, 3, 4, 8, and 9 (**Figure 4**). According to the reference genome sequence of “Hangzhou Gourd” (Wang et al., 2018), we preliminarily predicted three candidate genes in candidate regions or LD block regions of these 10 SNP markers (**Figure 5**). *HG_GLEAN_10011803*, a candidate gene, which was located 280 kb upstream of the BGRSe_00818 marker on Chr.1, encodes calcium-dependent protein kinase (CDPK) protein. There have been increasing studies confirming the involvement of CDPKs in plant disease resistance defense responses (Boudsocq and Sheen, 2013). For example, Loss-*AtCPK28* or overexpression-*AtCDPK1* mutants displayed enhanced responses to antibacterial immunity in *Arabidopsis* (Coca and Segundo, 2010; Monaghan et al., 2014). *SlCRK6* in tomato played a role in resistance to both *Sclerotinia sclerotiorum* and *Pseudomonas syringae* pv. *tomato* (Pst) DC3000 (Wang et al., 2016). *StCDPK5VK* in potato could increase resistance to late blight fungus through the production of ROS (Kobayashi et al., 2012). In addition, by conducting a qRT-PCR analysis, we found that the expression level of *HG_GLEAN_10011803* in FW-infected plants was significantly higher than that in healthy plants (**Figure 6**). Therefore, we inferred that the candidate gene *HG_GLEAN_10011803* might be related to the FW resistance of bottle gourd.

In the LD block region of another candidate marker BGRSe_14202, one candidate gene *HG_GLEAN_10001030*, located 50 kb upstream of this marker on chromosome 9, encoded the ethylene-responsive transcription factor (ERTF) RAP2 protein. ERTFs play an important regulatory role in plant signal transduction of disease resistance and stress resistance, and overexpression could improve plant disease resistance and stress resistance (Singh et al., 2002; Gutterson and Reuber, 2004). For example, *OsRAP2.6*-overexpressed plants showed improved resistance to rice blast fungus (Wamaita et al., 2012). *TERF1* and *TSRF1* genes in tomato could be resistant to *Ralstonia solanacearum* and *Botrytis cinerea* (Huang et al., 2004; Zhang et al., 2004, 2008). Another candidate gene, *HG_10001042*, located 18 kb downstream of this marker on chromosome 9, is a member of the GDSL gene family. The GDSL gene family consists of a wide range of members and plays important roles in plant growth, development, and stress defense responses (Akoh et al., 2004; Chepyshko et al., 2012). Overexpressed GDSL genes, such as *AtGLIP1* and *CaGLIP1*, could enhance the resistance to a variety of pathogenic fungi (Hong et al., 2008; Lee et al., 2009; Naranjo et al., 2010). The qPCR results showed that the expression levels of these two candidate genes were significantly increased in FW-infected YD-4 (HS material, level 4), while their expression levels were not significantly different before and after infection of Yin-10/Hanbi (HR materials, level 1) (**Figure 6**). Thus, we postulate that these three genes were candidate genes for FW resistance; in particular, *HG_GLEAN_10011803* might be a major effect gene. However, further evidence is needed to functionally validate this hypothesis. To our knowledge, this study is the first to perform GWAS for FW resistance in cucurbit crops. Our results provide the molecular tools for FW resistance selection and lay a foundation for candidate gene discovery.

The resistant materials and SNP markers that we identified will promote breeding programs for FW-resistant bottle gourd.

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

YL and GL conceived and designed the research. YL, YW, and JW performed the experiments. YL wrote the manuscript. YL, YW, and XinW constructed the population and collected phenotypes. XiaW, BW, and ZL carried out the field work. All authors analyzed the data and approved the submitted version.

FUNDING

This study was supported by the Special Sci-Tech Key Project of Breeding of New Varieties of Agricultural (Vegetables) Crops in Zhejiang Province (2016C02051-3-1), the Collaborative Extension Program of Agricultural Major Technology in Zhejiang Province (2019XTTGSC02-2), the Postdoctoral Research Program in Zhejiang Province and the Special Sci-Tech Key Project of Breeding of New Varieties of Vegetables in Zhejiang Province the 14th five-year plan.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.650157/full#supplementary-material>

Supplementary Figure 1 | Morphological features of wilted plant and *Fusarium oxysporum* f. sp. *lagenariae*. (A) Symptoms of wilted bottle gourd. (B) Symptoms of vascular bundle browning of wilted plants. (C) *F. oxysporum* colony of *Fusarium* wilt on potato dextrose agar. (D) Microscopic view of conidia of *F. oxysporum* f. sp. *lagenariae*.

Supplementary Figure 2 | Plant symptoms of *Fusarium* wilt disease at levels 0–4 in bottle gourd. From left to right: level 0 (I), level 1 (HR), level 2 (R), level 3 (S), and level 4 (HS).

Supplementary Figure 3 | Correlation analysis of disease index of 89 bottle gourd accessions between 2019 and 2020.

Supplementary Figure 4 | Q–Q plot of genome-wide association study for *Fusarium* wilt resistance based on four different association analysis models. (A) Four different association analysis models of DI2019. (B) Four different association analysis models of DI2020. Different colors represent different models: blue, GLM (PCA); black, GLM (Q); green, MLM (PCA+K); red, MLM (Q+K).

Supplementary Table 1 | Accession, origin, and disease index of 89 bottle gourd accessions used in this study.

Supplementary Table 2 | Primer sequences used for qRT-PCR.

Supplementary Table 3 | Significant markers associated with *Fusarium* wilt resistance in at least one environment.

REFERENCES

- Akroh, C. C., Lee, G. C., Liaw, Y. C., Huang, T. H., and Shaw, J. F. (2004). GDSL family of serine esterases/lipases. *Prog. Lipid Res.* 43, 534–552. doi: 10.1016/j.plipres.2004.09.002
- Anderson, A. D., and Weir, B. S. (2007). A maximum-likelihood method for the estimation of pairwise relatedness in structured populations. *Genetics* 176, 421–440. doi: 10.1534/genetics.106.063149
- Atwell, S., Huang, Y., Vilhjálmsson, B., Willems, G., Horton, M., Li, Y., et al. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465, 627–631. doi: 10.1038/nature08800
- Bhawna, A. M., Arya, L., Saha, D., Sureja, A. K., and Pandey, C. (2014). Population structure and genetic diversity in bottle gourd [*Lagenaria siceraria* (Mol.) Standl.] germplasm from India assessed by ISSR markers. *Plant Syst. Evol.* 300, 767–773. doi: 10.1007/s00606-014-1000-5
- Bodah, E. T. (2017). Root rot diseases in plants: a review of common causal agents and management strategies. *Agric. Res. Technol.* 5:555661. doi: 10.19080/ARTOAJ.2017.05.555661
- Boudsocq, M., and Sheen, J. (2013). CDPKs in immune and stress signaling. *Trends Plant Sci.* 18, 30–40. doi: 10.1016/j.tplants.2012.08.008
- Catanzariti, A. M., Lim, G. T., and Jones, D. A. (2015). The tomato I-3 gene: a novel gene for resistance to *Fusarium* wilt disease. *New Phytol.* 207, 106–118. doi: 10.1111/nph.13348
- Cha, J. Y., Han, S., Hong, H. J., Cho, H., Kim, D., Kwon, Y., et al. (2016). Microbial and biochemical basis of a *Fusarium* wilt-suppressive soil. *ISME J.* 10, 119–129. doi: 10.1038/ismej.2015.95
- Chepyshko, H., Lai, C. P., Huang, L. M., Liu, J. H., and Shaw, J. F. (2012). Multifunctionality and diversity of GDSL esterase/lipase gene family in rice (*Oryza sativa* L. japonica) genome: new insights from bioinformatics analysis. *BMC Genomics* 13:309. doi: 10.1186/1471-2164-13-309
- Coca, M., and Segundo, B. S. (2010). AtCPK1 calcium-dependent protein kinase mediates pathogen resistance in *Arabidopsis*. *Plant Mol. Biol.* 63, 526–540. doi: 10.1111/j.1365-3113X.2010.04255.x
- Ehlers, J. D., Hall, A. E., Patel, P. N., Roberts, P. A., and Matthews, W. C. (2000). Registration of ‘California blackeye 27’ cowpea. *Crop Sci.* 40, 854–855. doi: 10.2135/cropsci2000.403611x
- Ehlers, J. D., Sanden, B. L., Frate, C. A., Hall, A. E., and Roberts, P. A. (2009). Registration of ‘California blackeye 50’ cowpea. *J. Plant Regist.* 3, 236–240. doi: 10.3198/jpr2009.01.0039rc
- Erickson, D. L., Smith, B. D., Clarke, A. C., Sandweiss, D. H., and Tuross, N. (2005). An Asian origin for a 10,000-year-old domesticated plant in the Americas. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18315–18320. doi: 10.1073/pnas.0509279102
- 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
- Fang, C., Ma, Y. M., Wu, S. W., Liu, Z., Wang, Z., Yang, R., et al. (2017). Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. *Genome Biol.* 18:161. doi: 10.1186/s13059-017-1289-9
- Freeman, S., Zveibil, A., Vintal, H., and Maymon, M. (2002). Isolation of nonpathogenic mutants of *Fusarium oxysporum* f. sp. melonis for biological control of *Fusarium* wilt in cucurbits. *Phytopathology* 92, 164–168. doi: 10.1094/phyto.2002.92.2.164
- Gao, P., Liu, S., Zhu, Q. L., and Luan, F. S. (2015). Marker-assisted selection of *Fusarium* wilt-resistant and gynoecious melon (*Cucumis melo* L.). *Genet. Mol. Res.* 14, 16255–16264. doi: 10.4238/2015
- Guttersen, N., and Reuber, T. L. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr. Opin. Plant Biol.* 7, 465–471. doi: 10.1016/j.pbi.2004.04.007
- He, J., Meng, S., Zhao, T., Xing, G., and Gai, J. (2017). An innovative procedure of genome-wide association analysis fits studies on germplasm population and plant breeding. *Theor. Appl. Genet.* 130, 2327–2343. doi: 10.1007/s00122-017-2962-9
- Heiser, C. B. (1979). *Gourd Book*. Norman, OK: University of Oklahoma Press.
- Hong, J. K., Choi, H. W., Hwang, I. S., Kim, D. S., Kim, N. H., Choi, D. S., et al. (2008). Function of a novel GDSL-type pepper lipase gene, caglip1, in disease susceptibility and abiotic stress tolerance. *Planta* 227, 539–558. doi: 10.1007/s00425-007-0637-5
- Huang, X. H., Wei, X. H., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695
- Huang, Z., Zhang, Z., Zhang, X., Zhang, H., Huang, D., and Huang, R. (2004). Tomato TEF1 modulates ethylene response and enhances osmotic stress tolerance by activating expression of downstream genes. *FEBS. Lett.* 573, 110–116. doi: 10.1016/j.febslet.2004.07.064
- Joobeur, T., King, J. J., Nolin, S. J., Thomas, C. E., and Dean, R. A. (2004). The *Fusarium* wilt resistance locus Fom-2 of melon contains a single resistance gene with complex features. *Plant J.* 39, 283–297. doi: 10.1111/j.1365-313X.2004.02134.x
- Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., and Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics* 178, 1709–1723. doi: 10.1534/genetics.107.080101
- Katan, T. (1994). Physiologic races and vegetative compatibility groups of *Fusarium oxysporum* f. sp. melonis in Israel. *Phytopathology* 84, 153–157. doi: 10.1094/Phyto-84-153
- Khan, N., Maymon, M., and Hirsch, A. M. (2017). Combating *Fusarium* infection using *Bacillus*-based antimicrobials. *Microorganisms* 5:75. doi: 10.3390/microorganisms5040075
- Kobayashi, M., Yoshioka, M., Asai, S., Nomura, H., Kuchimura, K., Mori, H., et al. (2012). StCDPK5 confers resistance to late blight pathogen but increases susceptibility to early blight pathogen in potato via reactive oxygen species burst. *New Phytol.* 196, 223–237. doi: 10.1111/j.1469-8137.2012.04226.x
- Lambel, S., Lanini, B., Vivoda, E., Fauve, J., Patrick Wechter, W., Harris-Shultz, K. R., et al. (2014). A major QTL associated with *Fusarium oxysporum* race 1 resistance identified in genetic populations derived from closely related watermelon lines using selective genotyping and genotyping-by-sequencing for SNP discovery. *Theor. Appl. Genet.* 127, 2105–2115. doi: 10.1007/s00122-014-2363-2
- Lee, D. S., Kim, B. K., Kwon, S. J., Jin, H. C., and Park, O. K. (2009). *Arabidopsis* GDSL lipase 2 plays a role in pathogen defense via negative regulation of auxin signaling. *Biochem. Biophys. Res. Commun.* 379, 1038–1042. doi: 10.1016/j.bbrc.2009.01.006
- Lee, J. M. (1994). Cultivation of grafted vegetables I. current status, grafting methods, and benefits. *HortScience* 29, 235–239. doi: 10.21273/hortsci.29.4.235
- Li, H., Peng, Z., Yang, X., Wang, W., Fu, J., Wang, J., et al. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45, 43–50. doi: 10.1038/ng.2484
- Li, T., Ma, X., Li, N., Zhou, L., Liu, Z., and Han, H. (2017). Genome-wide association study discovered candidate genes of Verticillium wilt resistance in upland cotton (*Gossypium hirsutum* L.). *Plant Biotechnol. J.* 15, 1520–1532. doi: 10.1111/pbi.12734
- Li, Y., Reif, J. C., Ma, Y., Hong, H., Liu, Z., Chang, R., et al. (2015). Targeted association mapping demonstrating the complex molecular genetics of fatty acid formation in soybean. *BMC Genomics* 16:841. doi: 10.1186/s12864-015-2049-4
- 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
- Martyn, R. D., and Bruton, B. D. (1989). An initial survey of the United States for races of *Fusarium oxysporum* f. sp. *niveum*. *HortScience* 24, 696–698.
- Miguel, A., Maroto, J. V., Bautista, A. S., Baixauli, C., Cebolla, V., Pascual, B., et al. (2004). The grafting of triploid watermelon is an advantageous alternative to soil fumigation by methyl bromide for control of *Fusarium* wilt. *Sci. Hortic.* 103, 9–17. doi: 10.1016/j.scienta.2004.04.007
- Monaghan, J., Matschi, S., Shorinola, O., Rovenich, H., Matei, A., Segonzac, C., et al. (2014). The calcium-dependent protein kinase CPK28 buffers plant immunity and regulates BIK1 turnover. *Cell Host Microbe* 16, 605–615. doi: 10.1016/j.chom.2014.10.007
- Morimoto, Y., and Mvere, B. (2004). “*Lagenaria siceraria*,” in *Plant Resources of Tropical Africa* 2, eds G. J. H. Grubben, and O. A. Denton (Wageningen: Backhuys Publishers), 353–358.
- Naranjo, M. A., Forment, J., Roldán, M., Serrano, R., and Vicente, O. (2010). Overexpression of *Arabidopsis thaliana* LTL1, a salt-induced gene encoding a GDSL-motif lipase, increases salt tolerance in yeast and transgenic plants. *Plant Cell Environ.* 29, 1890–1900. doi: 10.1111/j.1365-3040.2006.01565.x

- Oumouloud, A., El-Otmani, M., Chikh-Rouhou, H., Garcés Claver, A., González Torres, R., Perl-Treves, R., et al. (2013). Breeding melon for resistance to *Fusarium* wilt: recent developments. *Euphytica* 192, 155–169. doi: 10.1007/s10681-013-0904-4
- Pottorff, M., Li, G., Ehlers, J. D., Close, T. J., and Roberts, P. A. (2013). Genetic mapping, synteny, and physical location of two loci for *Fusarium oxysporum* f. sp. *tracheiphilum* race 4 resistance in cowpea [*Vigna unguiculata* (L.) Walp]. *Mol. Breed.* 33, 779–791. doi: 10.1007/s11032-013-9991-0
- Pottorff, M., Wanamaker, S., Ma, Y. Q., Ehlers, J. D., Roberts, P. A., and Close, T. J. (2012). Genetic and physical mapping of candidate genes for resistance to *Fusarium oxysporum* f. sp. *tracheiphilum* race 3 in cowpea [*Vigna unguiculata* (L.) Walp]. *PLoS One* 7:e41600. doi: 10.1371/journal.pone.0041600
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Sabbavarapu, M. M., Sharma, M., Chamarthi, S. K., Swapna, N., Rathore, A., Thudi, M., et al. (2013). Molecular mapping of QTLs for resistance to *Fusarium* wilt (race 1) and *Ascochyta* blight in chickpea (*Cicer arietinum* L.). *Euphytica* 193, 121–133. doi: 10.1007/s10681-013-0959-2
- Shen, Z., and Li, X. X. (2008). *Descriptions and Data Standard for Bottle Gourd [Lagenaria siceraria (Mol.) Standl.]*. Beijing: China Agriculture Press, 56–58.
- Simons, G., Groenendijk, J., Wijbrandi, J., Reijans, M., Groenen, J., and Diegaarde, P. (1998). Dissection of the *Fusarium* 12 gene cluster in tomato reveals six homologs and one active gene copy. *Plant Cell* 10, 1055–1068. doi: 10.1105/tpc.10.6.1055
- Singh, K. B., Foley, R. C., and Luis, O. S. (2002). Transcription factors in plant defense and stress responses. *Curr. Opin. Plant Biol.* 5, 430–436. doi: 10.1016/S1369-5266(02)00289-3
- Singh, V. K., Singh, H. B., and Upadhyay, R. S. (2017). Role of fusaric acid in the development of 'Fusarium wilt' symptoms in tomato: physiological, biochemical and proteomic perspectives. *Plant Physiol. Biochem.* 118, 320–332. doi: 10.1016/j.plaphy.2017.06.028
- Wamaita, M. J., Yamamoto, R., Wong, H. L., Kawasaki, T., Kawano, Y., and Shimamoto, K. (2012). OsRap2.6 transcription factor contributes to rice innate immunity through its interaction with receptor for activated kinase-C1 (RACK1). *Rice* 5:35. doi: 10.1186/1939-8433-5-35
- Wang, J. P., Xu, Y. P., Munyampundu, J. P., Liu, T. Y., and Cai, X. Z. (2016). Calcium-dependent protein kinase (CDPK) and CDPK-related kinase (CRK) gene families in tomato: genome-wide identification and functional analyses in disease resistance. *Mol. Genet. Genomics* 291, 661–676. doi: 10.1007/s00438-015-1137-0
- Wang, P., Su, L., Qin, L., Hu, B., Guo, W., and Zhang, T. (2009). Identification and molecular mapping of a *Fusarium* wilt resistant gene in upland cotton. *Theor. Appl. Genet.* 119, 733–739. doi: 10.1007/s00122-009-1084-4
- Wang, Y., Xu, P., Wu, X. H., Wu, X. Y., Wang, B. G., Huang, Y. P., et al. (2018). GourdBase: a genome-centered multi-omics database for the bottle gourd (*Lagenaria siceraria*), an economically important cucurbit crop. *Sci. Rep.* 8:3604. doi: 10.1038/s41598-018-22007-3
- Wechter, W. P., Kousik, C., Mcmillan, M., and Leviet, A. (2012). Identification of resistance to *Fusarium oxysporum* f. sp. *niveum* race 2 in *Citrullus lanatus* var. *citroides* plant introductions. *HortScience* 47, 334–338. doi: 10.21273/HORTSCI.47.3.334
- Whitaker, T. W. (1971). *Endemism and Pre-Columbian Migration of Bottle Gourd, Lagenaria siceraria (Mol.) Standl.* Austin, TX: Man across the Sea, University of Press, 64–69.
- Wu, X., Xu, P., Wang, B., Lu, Z., and Li, G. (2017). Genome-wide association analysis of free glutamate content, a key factor conferring umami taste in the bottle gourd [*Lagenaria siceraria* (Mol.) standl.]. *Sci. Horticul.* 225, 795–801. doi: 10.1016/j.scienta.2017.08.015
- Xu, P., Wu, X., Luo, J., Wang, B., Liu, Y., and Ehlers, J. D. (2011). Partial sequencing of the bottle gourd genome reveals markers useful for phylogenetic analysis and breeding. *BMC Genomics* 12:467. doi: 10.1186/1471-2164-12-467
- Xu, P., Xu, S., Wu, X., Tao, Y., Wang, B., and Wang, S. (2014). Population genomic analyses from low-coverage RAD-Seq data: a case study on the non-model cucurbit bottle gourd. *Plant J.* 77, 430–442. doi: 10.1111/tpj.12370
- Xu, X. W., Yu, T., Xu, R., Shi, Y., Lin, X., Xu, Q., et al. (2016). Fine mapping of a dominantly inherited powdery mildew resistance major-effect QTL, Pm1.1, in cucumber identifies a 41.1kb region containing two tandemly arrayed cysteine-rich receptor-like protein kinase genes. *Theor. Appl. Genet.* 129, 507–516. doi: 10.1007/s00122-015-2644-4
- Yang, N., Lu, Y., Yang, X. H., Huang, J., Zhou, Y., and Ali, F. H. (2014). Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17 agronomic traits in an enlarged maize association panel. *PLoS Genet.* 10:e1004573. doi: 10.1371/journal.pgen.1004573
- Yetisir, H., and Sari, N. (2003). Effect of different rootstock on plant growth, yield and quality of watermelon. *Aust. J. Exp. Agric.* 43, 1269–1274. doi: 10.1071/EA02095
- Yu, J., Pressoir, G., Briggs, W. H., Vroh, B. I., 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
- Zhang, H., Yang, Y., Zhang, Z., Chen, J., Wang, X. C., and Huang, R. (2008). Expression of the ethylene response factor gene TSRF1 enhances abscisic acid responses during seedling development in tobacco. *Planta* 228, 777–787. doi: 10.1007/s00425-008-0779-0
- Zhang, H., Zhang, D., Chen, J., Yang, Y., Huang, Z., Huang, D., et al. (2004). Tomato stress-responsive factor TSRF1 interacts with ethylene responsive element GCC box and regulates pathogen resistance to *Ralstonia solanacearum*. *Plant Mol. Biol.* 55, 825–834. doi: 10.1007/s11103-004-2140-8
- Zhang, S. P., Miao, H., Yang, Y. H., Xie, B. Y., Wang, Y., and Gu, X. F. (2014). A major quantitative trait locus conferring resistance to *Fusarium* wilt was detected in cucumber by using recombinant inbred lines. *Mol. Breed.* 34, 1805–1815. doi: 10.1007/s11032-014-0140-1
- Zhang, X., Zhao, J., Bu, Y., Xue, D., Liu, Z., Li, X., et al. (2018). Genome-wide association studies of soybean seed hardness in the Chinese mini core collection. *Plant Mol. Biol. Rep.* 36, 605–617. doi: 10.1007/s11105-018-1102-2
- Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546
- Zhao, Y., Wang, H., Chen, W., and Li, Y. (2014). Genetic structure, linkage disequilibrium and association mapping of Verticillium wilt resistance in elite cotton (*Gossypium hirsutum* L.) germplasm population. *PLoS One* 9:e86308. doi: 10.1371/journal.pone.0086308
- Zink, F. W., and Gubler, W. D. (1985). Inheritance of resistance in muskmelon to *Fusarium* wilt. *J. Am. Soc. Hort.* 110, 600–604.

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 Li, Wang, Wu, Wang, Wu, Wang, Lu and Li. 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 and Mapping of Spot Blotch in *Triticum durum*–*Aegilops speltoides* Introgression Lines Using SNP Markers

OPEN ACCESS

Edited by:

Harsh Raman,
NSW Department of Primary
Industries, Australia

Reviewed by:

Elisabetta Mazzucotelli,
Council for Agricultural
and Economics Research (CREA),
Italy
Francesca Taranto,
Institute of Bioscience
and Bioresources, Italian National
Research Council, Italy

*Correspondence:

Parveen Chhuneja
pchhuneja@pau.edu

†ORCID:

Jaspal Kaur
orcid.org/0000-0002-5628-7325
Guriqbal Singh Dhillon
orcid.org/0000-0001-6766-810X
Satinder Kaur
orcid.org/0000-0003-3704-3074
Parveen Chhuneja
orcid.org/0000-0002-8599-9479

Specialty section:

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

Received: 07 January 2021

Accepted: 20 April 2021

Published: 28 May 2021

Citation:

Kaur J, Kaur J, Dhillon GS,
Kaur H, Singh J, Bala R, Srivastava P,
Kaur S, Sharma A and Chhuneja P
(2021) Characterization and Mapping
of Spot Blotch in *Triticum*
durum–*Aegilops speltoides*
Introgression Lines Using SNP
Markers. *Front. Plant Sci.* 12:650400.
doi: 10.3389/fpls.2021.650400

Jashanpreet Kaur¹, Jaspal Kaur^{2†}, Guriqbal Singh Dhillon^{3†}, Harmandeep Kaur¹,
Jasvir Singh², Ritu Bala², Puja Srivastava², Satinder Kaur^{3†}, Achla Sharma² and
Parveen Chhuneja^{3*†}

¹ Department of Plant Pathology, Punjab Agricultural University, Ludhiana, India, ² Department of Plant Breeding
and Genetics, Punjab Agricultural University, Ludhiana, India, ³ School of Agricultural Biotechnology, Punjab Agricultural
University, Ludhiana, India

Spot blotch (SB) of wheat is emerging as a major threat to successful wheat production in warm and humid areas of the world. SB, also called leaf blight, is caused by *Bipolaris sorokiniana*, and is responsible for high yield losses in Eastern Gangetic Plains Zone in India. More recently, SB is extending gradually toward cooler, traditional wheat-growing North-Western part of the country which is a major contributor to the national cereal basket. Deployment of resistant cultivars is considered as the most economical and ecologically sound measure to avoid losses due to this disease. In the present study, 89 backcross introgression lines (DSBILs) derived from *Triticum durum* (cv. PDW274-susceptible) × *Aegilops speltoides* (resistant) were evaluated against SB for four consecutive years, 2016–2020. Phenotypic evaluation of these lines showed a continuous variation in disease severity indicating that the resistance to SB is certainly quantitative in nature. Phenotypic data of DSBILs were further used for mapping QTLs using SNPs obtained by genotyping by sequencing. To identify QTLs stable across the environments, Best Linear Unbiased Estimates (BLUEs) and Predictions (BLUPs) were used for mapping QTLs based on stepwise regression-based Likelihood Ratio Test (RSTEP-LRT) for additive effect of markers and single marker analysis (SMA). Five QTLs, *Q.Sb.pau-2A*, *Q.Sb.pau-2B*, *Q.Sb.pau-3B*, *Q.Sb.pau-5B*, and *Q.Sb.pau-6A*, linked to SB resistance were mapped across chromosomes 2A, 2B, 3B, 5B, and 6A. Genes found adjacent to the SNP markers linked to these QTLs were literature mined to identify possible candidate genes by studying their role in plant pathogenesis. Further, highly resistant DSBIL (DSBIL-13) was selected to cross with a susceptible hexaploidy cultivar (HD3086) generating BC₂F₁ population. The QTL *Q.Sb.pau-5B*, linked to SNP S5B_703858864, was validated on this BC₂F₁ population and thus, may prove to be a potential diagnostic marker for SB resistance.

Keywords: leaf blight, spot blotch, backcross introgression lines, *Aegilops speltoides*, *Bipolaris sorokiniana*, *Triticum durum*, QTL

INTRODUCTION

Wheat, a major food crop of the world population, is in a constant threat from various biotic and abiotic stresses, limiting its potential for yield. Helminthosporium leaf blight/foiar blight/spot blotch (SB), caused by *Cochliobolus sativus* (anamorph: *Bipolaris sorokiniana*), is a major foliar disease of wheat in warmer wheat-growing regions. This hemibiotrophic fungus can potentially infect and damage various species of Poaceae family (Gupta et al., 2018). Due to drastic changes in the weather conditions in the last few decades leading to higher average temperature and unusual rainfall patterns, foliar leaf blight is emerging as a major threat to wheat production in India. Globally, an estimated 25 million ha of wheat land is affected by SB (Yadav et al., 2015), out of which around 10 million ha is in the Indian Subcontinent and 9 million ha of this is in the North-Eastern Plain Zone of India (Duveiller and Sharma, 2012; Chowdhury et al., 2013). This disease is extending gradually toward the North-West part characterized by high temperature and humidity late in the season (Saari, 1998) with an average yield loss of about 15–20% (Chand et al., 2003). The disease also causes serious damage in seed quality and market value of the produce leading to substantial economic losses (Singh and Kumar, 2008). Under heavy infestation, yield losses vary from 80 to 100% (Kumar et al., 2008). It is chiefly a seed-transmitted disease and the conidia can also survive in the soil.

Considering the huge wheat acreage attacked by this disease, it becomes obligatory to tackle this disease in wheat-growing areas through use of disease-free seed, seed treatment with a suitable fungicide reducing the carryover inoculum, and crop rotation to provide enough window period for decomposition of inoculum-carrying stubble (Chowdhury et al., 2013). Fungicide application seems to be the most convenient method. However, their repeated application involves significant cost, health hazard, and emergence of fungicidal resistance in the target pathogen. Among various alternatives, deployment of resistant cultivars remains a top priority approach as genetic resistance is an economical, robust, and environmentally friendly tool in the management of leaf blight disease. Resistance to leaf blight in the commonly grown wheat varieties of South-East Asia is generally insufficient or lacking (Joshi et al., 2004). So, there is an urgent need to identify sources of SB resistance from the gene pool of wild relatives.

From the limited number of inheritance studies, it has been found that both qualitative and quantitative type of inheritance are involved in SB resistance. A number of bi-parental studies and association mapping studies have reported QTLs linked to SB resistance present all over the wheat genome. Among them, four major QTLs, *Sb1* on 7D (Lillemo et al., 2013), *Sb2* on 5B (Kumar et al., 2015), *Sb3* on 3B (Lu et al., 2016), and *Sb4* on 4B (Zhang et al., 2020), have been identified and mapped. Several QTLs on chromosomes 2AL, 2BS, 5BL, and 6DL in “Yangmai#6”; on 2AS, 2BS, 5BL, and 7DS in the cultivar “Ning#8201”; and on 2BS, 2DS, 3BS, 7BS, and 7DS in the cultivar “Chirya#3” have been reported (Kumar et al., 2008, 2010). Neupane et al. (2010) reported a single, dominant gene conditioned resistance to leaf blight in

“Chirya#3” and “Milan/Sanghai#7.” Association mapping studies conducted by Gurung et al. (2014) and Adhikari et al. (2012) identified genomic regions associated with resistance to SB on chromosomes 1A, 1B, 3B, 5B, 6B, 7B, and 7D.

However, identification of donor lines resistant to SB remains a major continuing challenge (Joshi et al., 2007). At CIMMYT, a number of *Aegilops* and *Triticum* species were used as donors for resistance to leaf blight which included *Aegilops triuncialis*, *Aegilops cylindrica*, *Aegilops speltoides*, *Aegilops triaristata*, *Triticum dicoccoides* (wild emmer wheat), *Triticum boeoticum*, *Triticum persicum*, *Triticum timopheevii*, *Triticum araraticum*, *Triticum urartu*, and *Triticum sphaerococcum* (Singh and Dhaliwal, 1993; Smurova and Mikhailova, 2007). *Aegilops* species is considered as a good and less exploited source for increasing the genetic potential of cultivated wheat to various biotic and abiotic stresses.

In the wide hybridization program at Punjab Agricultural University, Ludhiana, a set of stable interspecific backcross introgression lines derived from *Triticum durum* and *A. speltoides* (DSBILs), putative B genome donor of wheat, were evaluated under polyhouse conditions for four consecutive seasons from 2016 to 2020 against SB. These DSBILs were used further to detect QTL(s) governing SB resistance and identify linked markers to aid in breeding for disease resistance in wheat. The linked markers were further used for validation on a BC₂F₁ population derived from HD3086 and one of the resistant DSBILs.

MATERIALS AND METHODS

Plant Genetic Material

A total of 89 backcross introgression lines derived from *A. speltoides* (accession #pau3809) and *T. durum* cultivar “PDW274” as recurrent parent were screened for resistance against SB. The F₁ plants from the cross of *T. durum* cv. PDW274 and *A. speltoides* acc. pau3809 were backcrossed for two generations with *T. durum* and selfed to generate BC₂F₁₀ introgression lines (DSBILs). Details of development of material can be retrieved from Awlchew et al. (2016).

Screening for SB Resistance

All the 89 DSBILs along with resistant parent *A. speltoides*, recurrent parent PDW274, and hexaploid susceptible check “Raj 4015” were evaluated under polyhouse conditions following artificially induced epiphytotic conditions. Susceptible check WL711 was sown after every 20 rows, and also in alleys to promote inoculum build-up and spread. Screening to leaf blight disease was done during four consecutive wheat seasons 2016–2017 (E1), 2017–2018 (E2), 2018–2019 (E3), and 2019–2020 (E4). Artificial epiphytotic conditions were created by spraying conidial suspension of the pathogen *B. sorokiniana* maintained on sorghum grains which were previously soaked and autoclaved. Aqueous conidial suspension (10⁶ conidia/ml) was sprayed on plants during evening hours until symptoms appeared on the susceptible checks. After inoculation, plants were lightly irrigated to provide high-humidity conditions, which is one

of the predisposing conditions for infection by *B. sorokiniana*. Disease scoring was done at three different growth stages (GS) on Zadoks' scale (Zadoks et al., 1974), which are GS55 (flowering stage or FS), GS75 (medium milk/dough stage or DS), and GS87 (hard dough stage or HDS), using a double-digit scale (00–99) which is based on percent leaf area covered due to blight in flag leaf and one leaf below flag leaf (F-1) as mentioned in **Supplementary Table 1**. The digit toward the left side indicates score of percent blighted area on flag leaf whereas the right digit gives the score of penultimate/F-1 leaf (Khan and Chowdhury, 2011). These two leaves at this stage contribute most to the grain-filling process thus directly affecting the grain yield (Chowdhury et al., 2013).

The AUDPC (area under disease progress curve) based on disease severity at GS55 (FS), GS75 (DS), and GS87 (HDS) was calculated as the total area under the graph of disease severity against time t , from the first disease evaluation to the last, with the following equation as given by Shaner and Finney (1977):

$$AUDPC = S_i = \sum_i^{n-1} [(t_{i+1} - t_i) (y_i + y_{i+1}) / 2]$$

where y_i = disease severity at time, $(t_{i+1} - t_i)$ = time in days between two disease scores, and n = number of dates for which SB disease level was recorded.

Statistical Analysis

The disease severity scores across different years using scores of FS, DS, HDS, and AUDPC were used to obtain best linear unbiased estimates (BLUEs) and predictions (BLUPs) by fitting linear mixed effects models in lme4 package v 1.1-26 (Bates et al., 2015) in R v4.0.3 (R Core Team, 2019) using

$$Y_{ik} = \mu + Year_i + Line_k + \varepsilon_{ik}$$

where Y_{ik} is the trait of interest, μ is the mean effect, $Year_i$ is the effect of the i th year, $Line_k$ is the effect of the k th line, and ε_{ik} is the error associated with the i th year and the k th line, which is assumed to be normally and independently distributed, with mean zero and homoscedastic variance σ^2 . For BLUEs model, the genotypes were considered as fixed effects, while for BLUPs model all the effects were considered as random effects. Considering genotypes as random effects reduces the effect of screening time along with other environmental effects on SB severity (Tomar et al., 2021). The disease severity scores obtained by fitting the BLUPs and BLUEs models were plotted using ggplot2 v3.3.3 (Wickham, 2016) and ggpvr v0.4.0 (Kassambara, 2020) in R v4.0.3 to study the distribution across the DSBILs.

Further, principal component analysis was performed to identify the number of principal components required to explain the variation across the years along with fitted values from linear mixed effect models using FactoMineR v2.4 (Lê et al., 2008) and factoextra v1.0.7 (Kassambara and Mundt, 2020) in R v4.0.3. The principal components were plotted as biplots to study the relation between disease severity scores of FS, DS, HDS, and AUDPC

along with identification of reduction in environmental effects in fitted values.

Genotyping

DNA extraction for all the 89 DSBILs along with both the parents was done using modified cetyltrimethylammonium bromide (CTAB) method (Saghai-Maroofo et al., 1984). All these DNA samples were genotyped with genotyping-by-sequencing (GBS) to provide dense genome-wide marker coverage. Raw sequence files were processed in the TASSEL GBS pipeline version 5.2.31 (Glaubitz et al., 2014) and further aligned to the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 reference genome. The vcf file so obtained was filtered for a minimum depth at 3 (DP3) and converted to hapmap format. The TASSEL output was then filtered for homozygous SNPs for each parental line and SNP markers polymorphic between the two parents were selected, and loci with very low coverage (<50%)/high missing data (>50%) or heterozygosity (>50%) were filtered out. DSBILs with more than 10% missing data were filtered out. Missing SNPs were imputed using the LD-kNNi method implemented in TASSEL with the following default parameters of minimum number of high LD sites = 30 and number of nearest neighbors = 10 (Ladejobi et al., 2019). SNPs with minor allele frequency (MAF) < 0.05 were excluded from further analysis and finally, 4056 SNPs with good quality genotype calls for 77 DSBILs along with recurrent parent were used for mapping.

QTL Mapping Using QTL IciMapping

QTL mapping was done by using CSL functionality of QTL IciMapping version 4.1 software (Meng et al., 2015). Disease resistance mapping was conducted with 4056 SNPs (MAF > 0.05) in 77 DS-BILs plus recurrent parent by stepwise regression-based Likelihood Ratio Test (RSTEP-LRT) for additive effect of markers and single marker analysis (SMA) in the software. Stepwise regression was used to determine the percentages of phenotypic variance explained (PVE) (R^2) by individual QTL and their respective additive effects at the likelihood of odds ratio (LOD) peaks. Significant SNPs were identified using threshold LOD of 3 at significant $p \leq 0.001$ and 1000 permutations. Only QTLs detected by both the algorithms and using both BLUPs and BLUEs were considered stable and significant. Further, the allelic effects were investigated to identify significantly associated markers with phenotypic data by Kruskal–Wallis test for studying the importance of individual alleles in SB disease resistance.

Postulation of Candidate Genes

The identified QTLs were further used to identify genic regions adjacent to their linked SNPs using reference genome assembly's functional annotation for high confidence genes (IWGSC Ref Seq v1.0). The genes were retrieved from a region of 500 kb on either side of the SNP and using the functional annotations, the proteins coded by these genes were identified. The functions of the proteins were further literature mined to identify their role in imparting resistance against SB.

TABLE 1 | Phenotypic evaluation for spot blotch disease severity of DSBILs along with recurrent parent (RP) and susceptible check.

| Env | Stage | RP | Check | Population | | | | | | |
|-------|-------|---------|---------|----------------|--------|--------|--------|------|-------|-------|
| | | PDW274 | Raj4015 | Range | Median | Mean | SD | CV | Skew. | Kurt. |
| BLUEs | FS | 14.50 | 38.00 | 00.25–39.00 | 11.25 | 12.71 | 8.42 | 0.66 | 0.96 | 0.54 |
| | DS | 62.00 | 70.25 | 08.75–70.50 | 42.50 | 41.02 | 13.83 | 0.34 | −0.20 | −0.86 |
| | HDS | 77.75 | 89.00 | 35.00–83.75 | 72.25 | 69.17 | 10.76 | 0.16 | −1.07 | 0.93 |
| | AUDPC | 1081.25 | 1337.50 | 302.50–1228.75 | 828.75 | 819.55 | 213.24 | 0.26 | −0.23 | −0.73 |
| BLUPs | FS | 13.61 | 24.20 | 07.18–24.65 | 12.14 | 12.81 | 3.79 | 0.30 | 0.96 | 0.56 |
| | DS | 52.92 | 57.58 | 22.82–57.73 | 41.90 | 41.09 | 7.76 | 0.19 | −0.21 | −0.83 |
| | HDS | 75.00 | 82.81 | 45.32–79.17 | 71.18 | 69.06 | 7.43 | 0.11 | −1.09 | 1.01 |
| | AUDPC | 985.25 | 1147.62 | 491.81–1078.71 | 825.26 | 819.86 | 134.52 | 0.16 | −0.24 | −0.70 |

RP, recurrent parent; CV, coefficient of variation; Skew., skewness; Kurt., kurtosis; Env, environment; BLUEs, best linear unbiased estimates; BLUPs, best linear unbiased predictions; FS, flowering stage; DS, dough stage; HDS, hard dough stage; AUDPC, area under disease progression curve.

Donor parent *Aegilops speltoides* (#pau3809) showed score 00 across all stages.

Validation of the Identified QTLs and Markers

For validation of the identified QTLs and markers, a BC₂F₁ population was developed from bread wheat cv. “HD3086” (high yielding, susceptible cultivar) and one of the DSBILs showing highly resistant response persistently under polyhouse conditions. All the plants were evaluated by creating artificial epiphytotic conditions as explained previously and scoring was done using a double-digit scale. Genomic DNA for all plants of BC₂F₁ mapping population and parents was extracted using CTAB method. To validate the SNP markers significantly linked to SB resistance as identified in mapping results, Kompetitive allele-specific PCR (KASP) assay was used for genotyping¹. For this purpose, SNPs linked to the QTLs were used to design KASP markers.

RESULTS

Phenotypic Evaluation of DSBILs

Large variation in disease severity was observed across the different growth stages with disease pressure increasing from flowering to hard dough stage (Table 1 and Supplementary Table 2). Across the environments, overall disease pressures were the lowest in E1 and highest in E3. To enhance the accuracy and map stable QTLs across the environments, linear mixed-effects models were used to obtain fitted values of disease severity, accounting for G × E effect. These values are termed as BLUPs (genotypes as random effects) and BLUEs (genotypes as fixed effects) from here onward. Plotting the eigenvalues/variances explained by each individual principal component (from PC1 to PC2), across the different growth stages and AUDPCs for all the environments, including BLUEs and BLUPs, showed that the first two principal components explained 93.7% of total (Supplementary Figure 1). The first two dimensions of principal components showed both BLUEs and BLUPs were able to explain the variance of disease scores across the four environments

(Figure 1). The BLUPs showed lower variance than the BLUEs which meant BLUPs were able to reduce the environmental variance across the years to a larger extent. The disease score distribution curves further agreed to this showing better normal distribution (Figure 2).

The donor parent *A. speltoides* acc. pau3809 was found to be immune to SB showing highly resistant disease severity score of 00, on the double-digit scale, across all the growth stages studied. Overall, the recurrent parent PDW274 showed moderate to high susceptibility and the susceptible check Raj4015 showed high susceptibility across the growth stages and AUDPCs when compared with range of disease scores of respective data sets. At FS, disease score BLUEs of the recurrent parent PDW274 and susceptible check Raj4015 were 14.5 and 38.0, respectively, while for the DSBILs, it ranged from 0.25 to 39.00. The disease score BLUPs of PDW274 and Raj4015 were 13.61, and 24.20, respectively with DSBILs showing a range from 07.18 to 24.65 (Table 1). At DS, disease score BLUEs of the recurrent parent PDW274 and susceptible check Raj4015 were 62.00 and 70.25, respectively, while for the DSBILs, it ranged from 08.75 to 70.50. The disease score BLUPs of PDW274 and Raj4015 were 52.92 and 57.58, respectively, with DSBILs showing a range from 22.82 to 57.73. At HDS, disease score BLUEs of the recurrent parent PDW274 and susceptible check Raj4015 were 77.75 and 89.00, respectively, while for the DSBILs, it ranged from 35.00 to 83.75. The disease score BLUPs of PDW274 and Raj4015 were 75.00 and 82.81, respectively, with DSBILs showing a range from 45.32 to 79.17.

The AUDPC values showed a similar trend, where the disease score BLUEs of the recurrent parent PDW274 and susceptible check Raj4015 were 1081.25 and 1337.50, respectively, while for the DSBILs, it ranged from 302.50 to 1228.75. The disease score BLUPs of PDW274 and Raj4015 were 985.25 and 1147.62, respectively, with DSBILs showing a range from 491.81 to 1078.71. Only three genotypes, DS13, DS61, and DS80, were found highly resistant across all the growth stages. Overall, less than 1% of lines were categorized under highly resistant category while 29 and 25% of genotypes showed moderate to high susceptibility, respectively. The rest of the lines fell under resistant to moderately resistant category.

¹<https://www.lgcgroup.com>

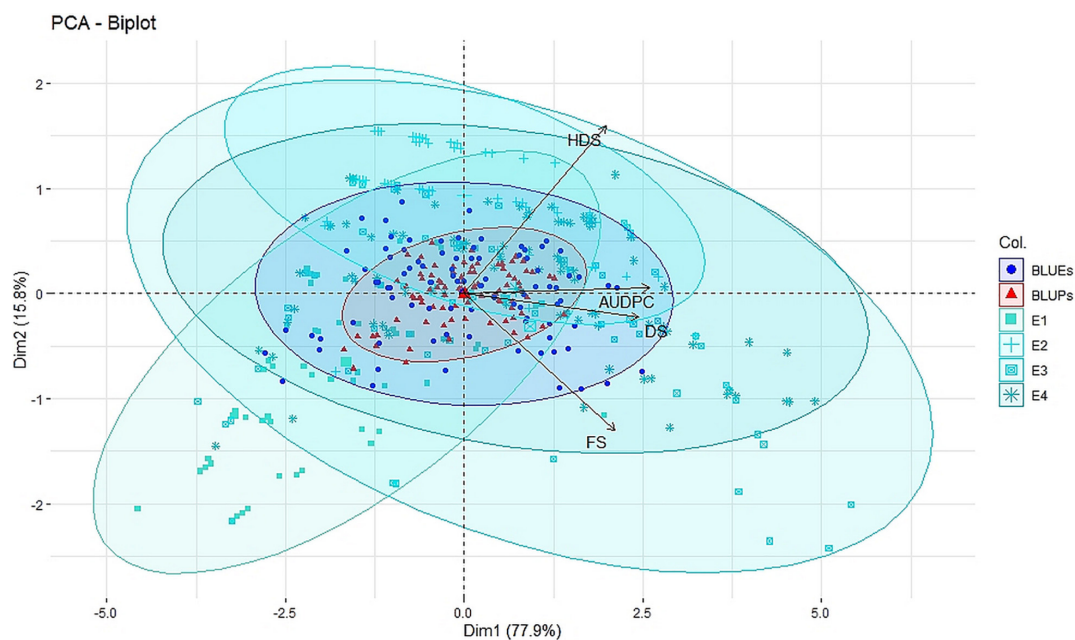


FIGURE 1 | Principal component analysis (PCA) biplot of disease severity of spot blotch across four individual environments E1, E2, E3, and E4 (turquoise color), across environment BLUEs (blue color), and across environment BLUPs (red color) for disease severity at flowering stage (FS), dough stage (DS), hard dough stage (HDS), and AUDPCs.

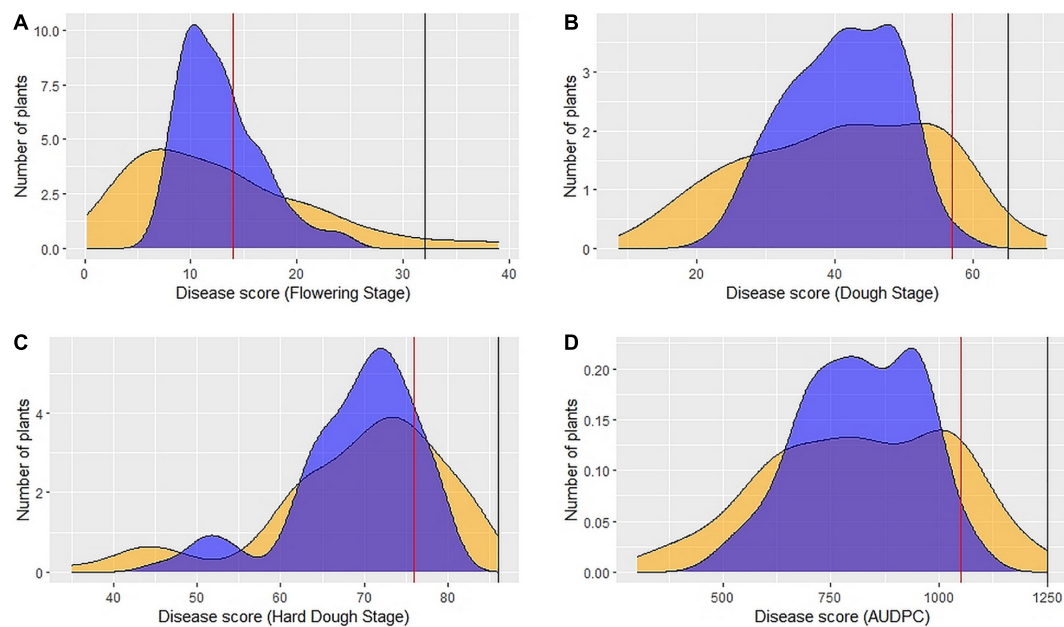


FIGURE 2 | Distribution of 89 DSBILs for spot blotch severity based on BLUEs (yellow color) and BLUPs (blue color) for disease score at flowering stage (A), dough stage (B), hard dough stage (C) and AUDPC (D). The vertical red line indicates disease score of recurrent parent PDW274, the vertical black line indicates disease score of susceptible check Raj4015, and the disease score of *A. speltoides* was 00 across the stages.

QTL Mapping

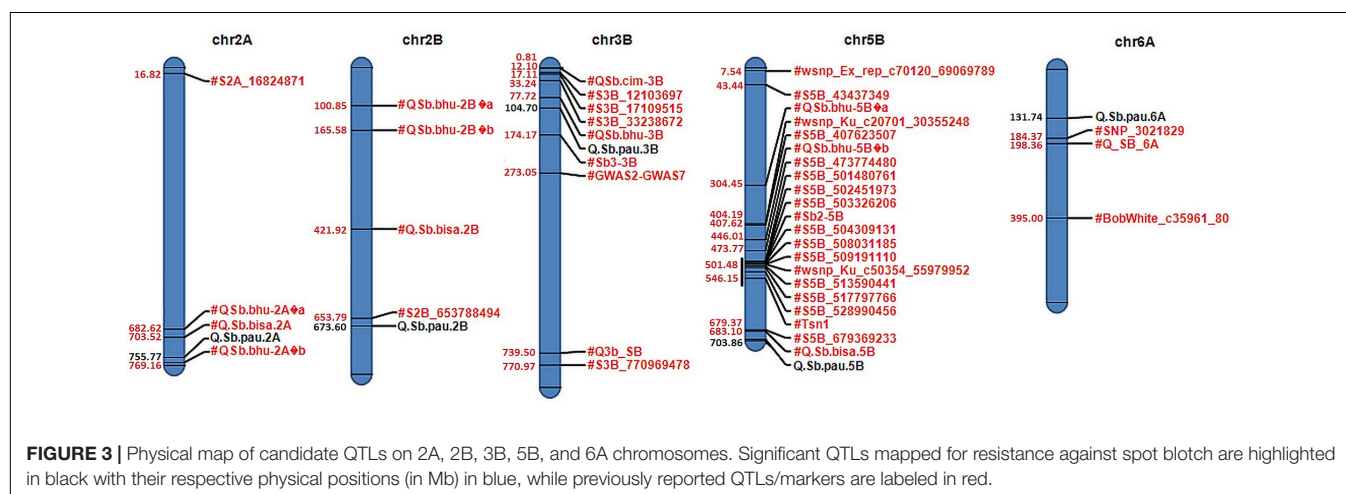
QTL mapping using both SMA and RSTEP-LRT for additive effect of markers using BLUPs and BLUEs for disease severity scores at different GS and AUDPCs resulted in detection of

five QTLs across five chromosomes (Table 2 and Figure 3). These QTLs were located on chromosomes 2A, 2B, 3B, 5B, and 6A. The phenotypic variation explained by these QTLs varied from 16.03 to 25.56%, while the LOD score varied from 3.04

TABLE 2 | Summary of the QTLs detected using both single marker analysis (SMA) and RSTEP-LRT for additive effect of markers algorithms of QTL ICI mapping for spot blotch disease severity.

| QTL | Marker | Chr | Pos (Mb) | GS | Env | LOD | PVE (%) | Add |
|--------------------|---------------|-----|----------|-------|-------|------|---------|---------|
| <i>Q.Sb.pau-2A</i> | S2A_755774702 | 2A | 755.77 | HDS | BLUEs | 3.12 | 18.44 | -8.24 |
| | | | | | BLUPs | 3.18 | 18.77 | -5.73 |
| <i>Q.Sb.pau-2B</i> | S2B_673595704 | 2B | 673.60 | AUDPC | BLUEs | 3.04 | 16.03 | 160.96 |
| | | | | | BLUPs | 3.16 | 16.98 | 149.57 |
| | | | | HDS | BLUEs | 4.09 | 21.27 | 11.81 |
| | | | | | BLUPs | 5.02 | 25.56 | 10.35 |
| <i>Q.Sb.pau-3B</i> | S3B_104700872 | 3B | 104.70 | AUDPC | BLUEs | 4.33 | 23.27 | -128.66 |
| | | | | | BLUPs | 4.58 | 25.27 | -113.13 |
| | | | | HDS | BLUEs | 3.22 | 17.86 | -8.62 |
| | | | | | BLUPs | 3.82 | 20.43 | -7.79 |
| <i>Q.Sb.pau-5B</i> | S5B_703858864 | 5B | 703.86 | DS | BLUEs | 3.32 | 19.48 | -6.06 |
| | | | | | BLUPs | 3.23 | 18.97 | -3.36 |
| <i>Q.Sb.pau-6A</i> | S6A_131743987 | 6A | 131.74 | FS | BLUEs | 3.07 | 16.38 | 7.59 |
| | | | | | BLUPs | 3.08 | 16.42 | 3.42 |

Chr, chromosome; Pos Mb, position in million bases; GS, growth stage; Env, environment; LOD, logarithm of odds; PVE, phenotypic variation explained; Add, additive effect; BLUEs, best linear unbiased estimates; BLUPs, best linear unbiased predictions; FS, flowering stage; DS, dough stage; HDS, hard dough stage; AUDPC, area under disease progression curve.

**FIGURE 3 |** Physical map of candidate QTLs on 2A, 2B, 3B, 5B, and 6A chromosomes. Significant QTLs mapped for resistance against spot blotch are highlighted in black with their respective physical positions (in Mb) in blue, while previously reported QTLs/markers are labeled in red.

to 5.02. QTL *Q.Sb.pau-2A* was mapped at chromosome 2A at 755.77 Mb using disease severity at hard dough stage with LOD 3.12, PVE 18.44% using BLUEs and LOD 3.18, PVE 18.77% using BLUPs with resistance allele contributed by *A. speltoides*. Two QTLs, *Q.Sb.pau-2B* and *Q.Sb.pau-3B*, were mapped using both disease severity at HDS and AUDPC, where the resistant allele for *Q.Sb.pau-2B* was contributed by PDW274 while the resistant allele for *Q.Sb.pau-3B* was contributed by *A. speltoides* in both cases. QTL *Q.Sb.pau-2B* was mapped at chromosome 2B at 673.60 Mb using disease severity at HDS with LOD 4.09, PVE 21.27% using BLUEs and LOD 5.02, PVE 25.56% using BLUPs. Using AUDPCs, it was mapped with LOD 3.04, PVE 16.03% using BLUEs and LOD 3.16, PVE 16.98% using BLUPs. QTL *Q.Sb.pau-3B* was mapped at chromosome 3B at 104.70 Mb using disease severity at HDS with LOD 3.22, PVE 17.86% using BLUEs and LOD 3.82, PVE 20.43% using BLUPs. Using AUDPCs, it was mapped with LOD 4.33, PVE 23.27% using BLUEs and LOD 4.58, PVE 25.27% using BLUPs.

Only one QTL was mapped for disease severity at FS and DS. QTL *Q.Sb.pau-5B* was mapped at chromosome 5B at 703.86 Mb using disease severity at DS with LOD 3.32, PVE 19.48% using BLUEs and LOD 3.23, PVE 18.97% using BLUPs with resistance allele contributed by *A. speltoides*. QTL *Q.Sb.pau-6A* was mapped at chr6A at 131.74 Mb using disease severity at FS with LOD 3.07, PVE 16.38% using BLUEs and LOD 3.08, PVE 16.42% using BLUPs with resistance allele contributed by PBW274.

Allelic Effect of Identified QTLs

The allelic effect of the SNPs linked to SB QTL was plotted for the five significant QTLs (Figure 4). The allelic variation patterns of the QTLs, between the two alternate alleles, further agreed with positive mapping results. Also, the donor parent of resistant alleles was confirmed for the QTLs. The disease severity scores of growth stages in which the QTLs were detected were used along with AUPDCs. The patterns of disease severity for

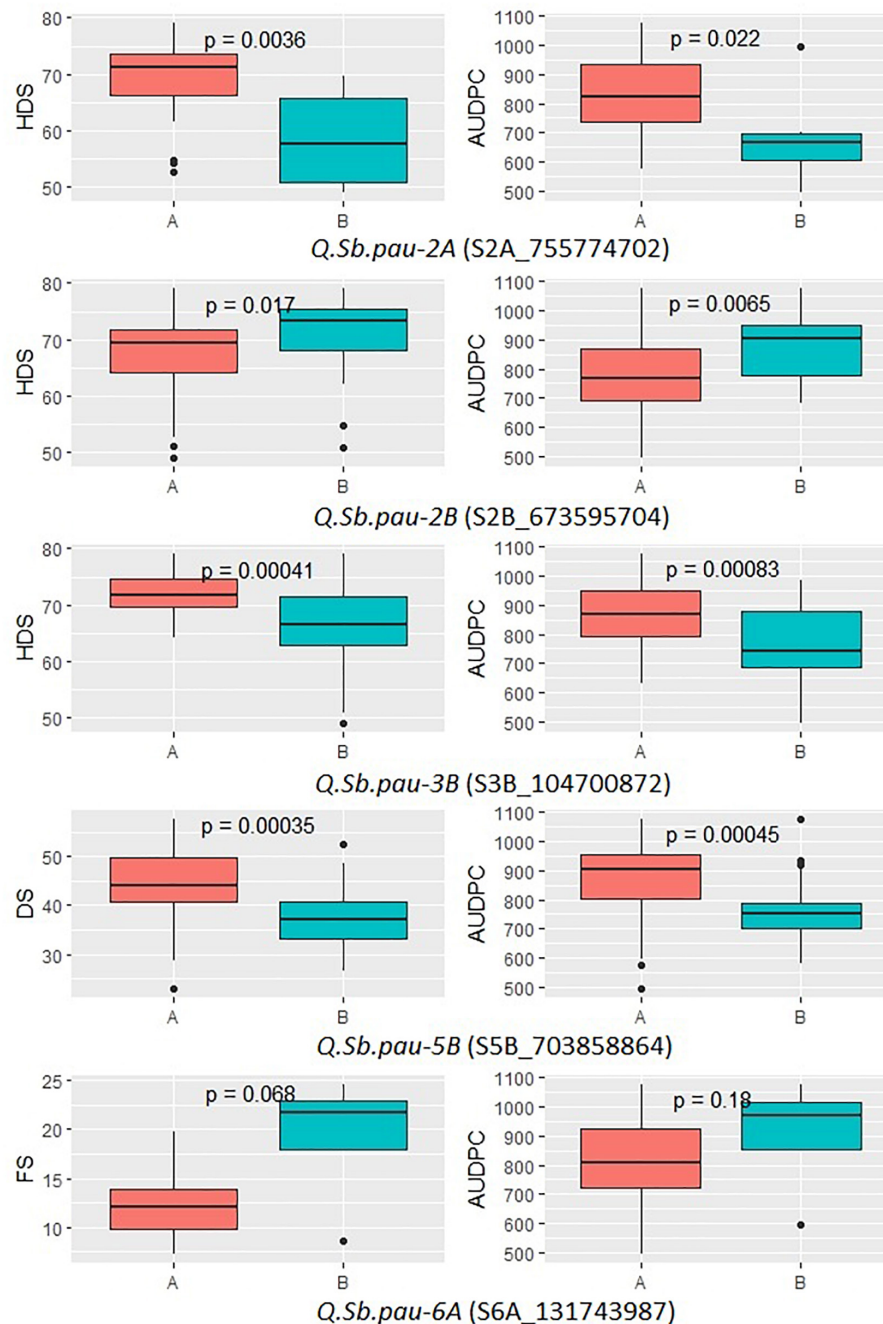


FIGURE 4 | Boxplots showing the effect of phenotypic variation between the two alleles of the SNP markers linked to QTLs for disease score of DSBILs. Kruskal–Wallis test was used to determine the significant differences between the mean values of two alleles.

the alternate alleles confirmed that three of the QTLs (*Q.Sb.pau-2A*, *Q.Sb.pau-3B*, and *Q.Sb.pau-5B*) had resistant allele donated by *A. speltoides*, while the remaining two QTLs (*Q.Sb.pau-2B*, and *Q.Sb.pau-6A*) had resistant allele donated by PDW274. The alternate alleles of QTL *Q.Sb.pau-3B* and *Q.Sb.pau-5B* were the most significantly different for respective disease severity score and AUDPCs, while the alternate alleles of QTL *Q.Sb.pau-6A* were least significantly different.

Postulation of Candidate Genes

Because the physical locations of the SNPs linked to the QTLs detected in the present study were known, they were used to identify the genes present adjacent to them in a region of 500 kb on either side of the SNP (**Supplementary Table 3**). For each target locus, the regions were inspected to identify candidate genes for the QTL and the genes known to be involved in different pathways of pathogen–host

TABLE 3 | QTLs along with SNPs and corresponding proteins and functional gene annotation elucidated based on the high confidence genes from wheat reference sequence (RefSeq V1.0) annotation database.

| QTL | SNP | Chr | GeneID | Dist. (kb) | Gene annotation |
|-------------|---------------|-----|---------------------------|------------|--|
| Q.Sb.pau-2A | S2A_755774702 | 2A | <i>TraesCS2A01G546600</i> | 439.059 | Cytochrome P450 family protein, expressed |
| | | | <i>TraesCS2A01G546700</i> | 435.192 | Cysteine proteinase |
| | | | <i>TraesCS2A01G546800</i> | 418.040 | Zinc finger MYM-type-like protein |
| | | | <i>TraesCS2A01G547400</i> | 81.978 | FBD, F-box and Leucine Rich Repeat domains protein |
| | | | <i>TraesCS2A01G547600</i> | 37.854 | Cytochrome P450, putative |
| | | | <i>TraesCS2A01G547800</i> | 5.678 | Auxin response factor |
| | | | <i>TraesCS2A01G547900</i> | -11.024 | Zinc finger CCCH domain-containing protein 32 |
| Q.Sb.pau-2B | S2B_673595704 | 2B | <i>TraesCS2B01G476400</i> | 425.139 | Senescence-associated family protein (DUF581) |
| | | | <i>TraesCS2B01G476500</i> | 353.761 | Senescence-associated family protein (DUF581) |
| | | | <i>TraesCS2B01G476600</i> | -32.658 | Senescence-associated family protein (DUF581) |
| | | | <i>TraesCS2B01G476700</i> | -112.168 | Senescence-associated family protein (DUF581) |
| | | | <i>TraesCS2B01G476800</i> | -115.776 | Senescence-associated family protein (DUF581) |
| | | | <i>TraesCS2B01G476900</i> | -329.070 | Senescence-associated family protein (DUF581) |
| | | | | | |
| Q.Sb.pau-3B | S3B_104700839 | 3B | <i>TraesCS3B01G127000</i> | 375.807 | Protein FAR1-RELATED SEQUENCE 3 |
| | | | <i>TraesCS3B01G127100</i> | -269.633 | IQ domain-containing protein |
| Q.Sb.pau-5B | S5B_703858864 | 5B | <i>TraesCS5B01G553200</i> | 370.394 | F-box family protein |
| | | | <i>TraesCS5B01G553300</i> | 361.335 | F-box domain containing protein |
| | | | <i>TraesCS5B01G553400</i> | 358.507 | F-box and associated interaction domains protein |
| | | | <i>TraesCS5B01G553500</i> | 329.318 | F-box domain containing protein, expressed |
| | | | <i>TraesCS5B01G553700</i> | 214.322 | F-box family protein |
| | | | <i>TraesCS5B01G553900</i> | 1.689 | F-box family protein |
| | | | <i>TraesCS5B01G554000</i> | -214.708 | ATP-dependent Clp protease ATP-binding subunit |
| | | | <i>TraesCS5B01G554100</i> | -232.540 | F-box family protein |
| | | | <i>TraesCS5B01G554200</i> | -250.728 | Disease resistance protein RPM1 |
| Q.Sb.pau-6A | S6A_131743987 | 6A | <i>TraesCS5B01G554300</i> | -276.340 | Disease resistance protein (NBS-LRR class) family |
| | | | <i>TraesCS5B01G554500</i> | -368.461 | AIG2-like (Avirulence induced gene) family protein |
| | | | <i>TraesCS6A01G149500</i> | 297.870 | Ubiquitin family protein |
| | | | <i>TraesCS6A01G149600</i> | -436.076 | Uricase |

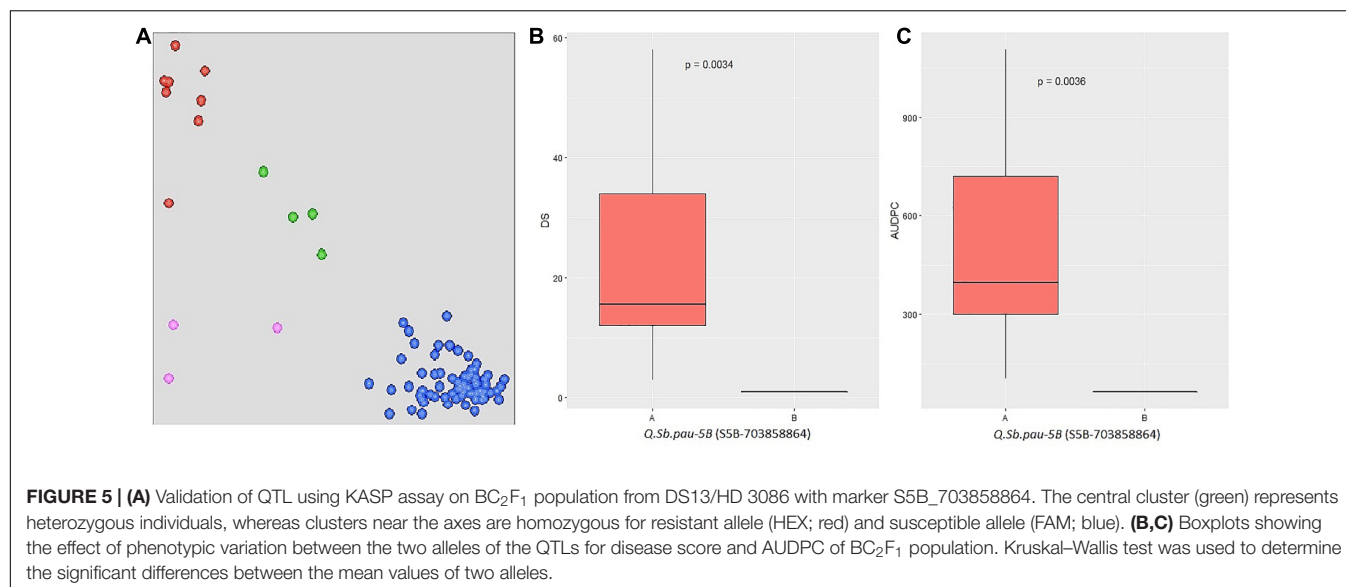
Distance from SNP (Dist.) represents distance of start site of gene to SNP linked with QTL, where (+) sign indicates that the gene was found downstream of the SNP and (-) sign indicates that the gene was found upstream.

interactions and pathogenesis were considered to understand their role in imparting resistance to SB (Table 3). The SNP S2A_755774702 linked to QTL Q.Sb.pau-2A was found adjacent to genes *TraesCS2A01G547800* and *TraesCS2A01G547900*. The gene *TraesCS2A01G547800* codes for Auxin response factor (ARF) and *TraesCS2A01G547900* codes for Zinc finger CCCH domain-containing protein 32. Gene *TraesCS2A01G547400* was found in close vicinity of the QTL for FBD, F-box, and Leucine Rich Repeat domains protein. Other genes in the genomic region with probable role in disease resistance coded for cysteine proteinase (*TraesCS2A01G546700*), Cytochrome P450 (*TraesCS2A01G546600* and *TraesCS2A01G547600*), and Zinc finger MYM-type-like protein (*TraesCS2A01G546800*). The SNP S2B_673595704 linked to QTL Q.Sb.pau-2B was found adjacent to genes *TraesCS2B01G476500* and *TraesCS2B01G476600* both encoding senescence-associated family protein (DUF581). Four other genes coding for DUF581 were also found in the genomic region of the QTL. The SNP S3B_104700839 linked to QTL Q.Sb.pau-3B was found adjacent to genes *TraesCS3B01G127000* and *TraesCS3B01G127100*. The gene *TraesCS3B01G127000* coded for protein FAR1-RELATED SEQUENCE 3 and gene *TraesCS3B01G127000* coded for IQ domain-containing protein.

The SNP S5B_703858864 linked to QTL Q.Sb.pau-5B was found adjacent to genes *TraesCS5B01G553900* and *TraesCS5B01G554000*. The gene *TraesCS5B01G553900* coded for F-box family protein and gene *TraesCS5B01G554000* coded for ATP-dependent Clp protease ATP-binding subunit. Along with six other F-box family protein coding genes in the region, three disease resistance protein genes RPM1 (*TraesCS5B01G554100*), NBS-LRR family protein (*TraesCS5B01G554200*), and AIG2 like protein (*TraesCS5B01G554300*) were found in the genomic region of the QTL. The SNP S6A_131743987 linked to QTL Q.Sb.pau-6A was found adjacent to genes *TraesCS6A01G149500* and *TraesCS6A01G149600*. The gene *TraesCS6A01G149500* coded for Ubiquitin family protein and gene *TraesCS6A01G149600* coded for uricase.

Validation of the Identified QTLs and Markers

BC₂F₁ population derived from DSBIL13 × HD3086 was generated for transfer of SB resistance into wheat and to validate the identified SNP markers linked to SB QTLs, where HD3086 is a high-yielding SB-susceptible hexaploid cultivar and DSBIL13



is a highly resistant line. Besides being highly resistant to SB, DSBIL13 also harbored resistant alleles of four out of five QTLs mapped in the present study, namely *Qsb.pau-2A*, *Qsb.pau-2B*, *Qsb.pau-3B*, and *Qsb.pau-5B*. About 75% of the plants were found to show resistance reaction when screened phenotypically under polyhouse conditions.

The SNPs linked to the five SB resistance QTLs were converted to KASP markers (Supplementary Table 4) and parental polymorphism survey was done to study the allelic composition of HD3086, DSBIL13, PDW274, and *A. speltoides* acc. pau3809. Out of five markers, only S5B_703858864 was found to be polymorphic between HD3086 and DSBIL13, i.e., HD3086 harbored an alternate allele to the allele imparting resistance. Thus, only this marker could be successfully used to track the SB resistance allele of QTL *Qsb.pau-5B*. This marker was then applied to BC₂F₁ population derived from DSBIL13 × HD3086. The disease severity scores of growth stages in which the QTL was detected was used along with AUPDCs to evaluate significance of differences by Kruskal test of significant difference (Figure 5). The patterns of disease severity for the alternate alleles confirmed that the QTL *Qsb.pau-5B* having resistant allele from *A. speltoides* was transferred to the BC₂F₁ population with significant difference of alternate alleles at $p = 0.0034$ for DS and $p = 0.0036$ for AUDPC. Thus, this marker can be used for marker-assisted selection (MAS) and gene pyramiding in future crop improvement programs.

DISCUSSION

Spot blotch is one of the major constraints to the global wheat production, especially in areas with hot and humid climate (Tomar et al., 2021). To counter the constraints from foliar diseases like SB, there is a need for constantly identifying and introgressing new sources of resistance. The DSBIL panel used

in the present study showed wide range of variation for different traits and has already been reported to possess various QTLs for heat tolerance, stripe rust, and powdery mildew resistance (Awlachev et al., 2016; Dhillon et al., 2020). In the present study, during phenotypic evaluation of disease severity for SB, three DSBILs (DSBIL 13, 61, and 80) were identified to be highly resistant against SB. Because no wheat cultivar presently grown in North-Western plains of India possess resistance to SB, these lines become an important resource for transfer of SB resistance. The phenotypic evaluation showed increased disease severity from FS (GS55) to HDS (GD87). At HDS, immunity was mostly characterized by moderate resistance in the DSBIL panel. The continuous distribution of disease severity score in the panel indicated additive effect leading to quantitative nature of resistance. Most of the studies on SB resistance dictate that multiple genes with minor effect control the SB resistance in wheat (Adhikari et al., 2012; Lillemo et al., 2013; Zhuang et al., 2013; Gurung et al., 2014; Lu et al., 2016; Ayana et al., 2018; Kaur et al., 2018; Singh et al., 2018; Tomar et al., 2021). Kumar et al. (2008) also suggested that the resistance to SB is polygenic and controlled by a number of loci each having its own additive effect. This was further confirmed by Singh et al. (2018) in which lines were continuously distributed based on phenotypic screening, indicating that the resistance to leaf blight is certainly quantitative in nature. Latwal et al. (2016) reported that out of 200 wheat accessions obtained from CIMMYT, ~5 lines were highly resistant, ~123 lines resistant to moderately resistant, and 28 lines were susceptible to highly susceptible. A similar pattern was observed in the present study.

In the present study, five QTLs were mapped on chr 2A, 2B, 3B, 5B, and 6A. Three out of five QTLs (*Qsb.pau-2A*, *Qsb.pau-3B*, and *Qsb.pau-5B*) had resistance allele donated by *A. speltoides* and are probably novel as no gene/QTL for SB resistance from *A. speltoides* has yet been reported, despite large genetic potential against SB (Smurova and Mikhailova, 2007). FBD, F-box, and

Leucine Rich Repeat domains protein was 80.98 kb from the QTL *QSh.pau-2A*. NBS-LRR genes are the most common disease resistance gene family in plant genomes (Lee and Yeom, 2015; Dubey and Singh, 2018). A gene coding for ARF protein was found 5.68 kb from the QTL. As reported by Fu and Wang (2011), ARF regulates (enhance or repress) the transcription of primary auxin-responsive genes, thus involving auxin in biotic stress defense responses. Auxin-responsive genes are downregulated in *Arabidopsis thaliana* upon *Botrytis cinerea* infection making it more susceptible (Llorente et al., 2008). The genomic region of this SNP harbored another gene with Zinc finger CCCH domain-containing protein 32 (AtC3H32). Tandem CCCH zinc finger (TZF) motifs are known to play a variety of roles: ABA and gibberellin stress response (Lin et al., 2011), seed germination (Kim et al., 2008), mediated pathogen-associated molecular pattern (PAMP)-triggered immune responses (Maldonado-Bonilla et al., 2013), and involved in salt stress responses (Sun et al., 2007). Maldonado-Bonilla et al. (2013) reported that in *A. thaliana*, tandem zinc finger protein is phosphorylated by PAMP-responsive MAPKs which is required to trigger a PAMP-triggered immunity (PTI). Two genes coding for Cytochrome P450 were also found in the genomic region of the QTL and wheat Cytochrome P450 family protein is known to induce resistance to mycotoxin deoxynivalenol (DON) (Gunupuru et al., 2018). The cysteine protease coding gene in the region of the QTL also plays an important role as the extracellular cysteine protease is important for pathogen recognition. An oxidative burst is triggered by recognition, accompanied by transcriptional reprogramming and HR, which leads to disease resistance (Thomas and van der Hoorn, 2018). QTL *QSh.pau-2A* mapped in the present study was found 55 Mb from QTL *QSh.bisa.2A* (Tomar et al., 2021) and in same genomic region of QTL *QSh.bhu-2A* (Kumar et al., 2010) and is probably novel as this QTL had been contributed by *A. speltoides* while previously reported QTL are from cultivated wheat.

QTL *QSh.pau-2B* was mapped 20 Mb from another QTL previously mapped in the region (Bainsla et al., 2020) and was found to be flanked by six senescence-associated family protein (DUF581) coding genes. In wheat, if one allele of the gene is involved in senescence, the other is associated with the stay-green trait (Tomar et al., 2021) and the stay-green trait has been reported to positively correlate with wheat leaf blight resistance (Joshi et al., 2006; Rosyara et al., 2008). QTL *QSh.pau-3B* was found to be linked with Protein FAR1-RELATED SEQUENCE 3 which is known to modulate plant immunity. FHY3 and its homolog FAR1 improve resistance by negatively regulating ROS accumulation and suppressing plant cell death (Ma and Li, 2018) and by positively regulating the biosynthesis of myo-inositol (Ma et al., 2016). The genomic region of the QTL was found to carry another gene coding for IQ domain-containing protein. Levy et al. (2005) reported that in *A. thaliana*, this protein, IQD1, encodes a novel nuclear protein that binds to calmodulin in a Ca^{2+} -dependent fashion and stimulates accumulation of plant defense-related secondary metabolite glucosinolates. QTL *QSh.pau-3B* was mapped 28 Mb from QTL *QSh.bhu-3B*

(Kumar et al., 2010) and hence this QTL introgressed from *A. speltoides* might be novel.

QTL *QSh.pau-5B* was found 20 Mb from earlier reported QTL *QSh.bisa.5B* (Tomar et al., 2021) and 24 Mb from another QTL *S5B_679369233* (Jamil et al., 2018). Chromosome 5B has been reported as hotspot for SB resistance as a large number of QTLs/genes mapped for resistance against SB have been mapped on this chromosome. The annotation study revealed that the SNP *S5B_703858864* linked to loci *QSh.pau-5B* is associated with three disease resistance protein coding genes. Both RP1 and AIG2 protein are known to play a crucial role in recognition of pathogens and effector-triggered immune responses in plants (Reuber and Ausubel, 1996; Beth Mudgett, 2005; Chisholm et al., 2006). The third resistance gene was NB-LRR gene which are the most common disease resistance gene family in plant genomes (Lee and Yeom, 2015; Dubey and Singh, 2018). The region also included six F-box family proteins. F-box family protein mediates a variety of biological processes, such as leaf senescence (Woo et al., 2001), and responses to biotic (Kim and Delaney, 2002) and abiotic stresses (Calderón-Villalobos et al., 2007). In mutant seedlings of *Arabidopsis* showing high susceptibility to pathogen *Peronospora parasitica*, Kim and Delaney (2002) have reported to isolate *son1* protein which was responsible to induce resistance among the seedlings. On cloning *son1*, it was found to encode a novel protein containing F-box motif, an element found within the E3 ubiquitin-ligase complex, suggesting the existence of a novel defense response through the ubiquitin-proteasome pathway, independent of SAR. The genomic region also carries gene encoding for ATP-dependent Clp protease ATP-binding subunit. Clp protease degrades damaged or non-native proteins in mitochondria and chloroplasts whose amount increases during abiotic and biotic stress conditions (Ali and Baek, 2020).

QTL *QSh.pau-6A* was mapped 53 Mb from QTL *SNP_3021829* (Bainsla et al., 2020) mapped in the same genomic region. A gene for Ubiquitin family protein was found flanking the QTL. Ubiquitin-related proteins impart plant resistance by degrading flagellin-sensing 2 (FLS2) receptor, which binds the microbe-associated molecular pattern (MAMP), flagellin (Trujillo and Shirasu, 2010; Lu et al., 2011). Ubiquitin, which is a part of the ubiquitin-proteasome system (UPS), controls various pathways including response to biotic and abiotic stresses (Sadanandom et al., 2012), and acts as one of the major systems in plant immunity (Üstün et al., 2016). Besides immunity, their role in defense responses by the production of ROS and forming hypersensitive reactions have also been reported (Marino et al., 2012). Another gene flanking the QTL coded for uricase. Increased activity of uricase has been observed in bean leaf tissue after infection with *Uromyces phaseoli* (Montalbini, 1991) in both resistant and susceptible plants. Higher activity of uricase was observed more in plants with hypersensitive reaction than in the susceptible plants.

The SNPs linked to QTLs were used to design KASP-based markers for marker-assisted transfer and validation. Using a susceptible high-yielding cultivar HD3086 and highly resistant DSBIL13, a BC₂F₁ population was generated. Since four of the five markers were not polymorphic between HD3086 and

DSBIL13, only one marker S5B_703858864 linked to QTL *Qsb.pau-5B* could be validated on the segregating population. The homozygous alternate alleles of this marker showed significant difference for SB severity with p value < 0.01 , and thus this marker could be used for marker-assisted transfer of the QTL. The phenotypic evaluation of the segregating population showed a wide range of SB severity scores from highly resistant to highly susceptible, which indicated that more than one locus for resistance was segregating in the population. This segregation pattern was highly expected as DSBIL13 harbored four QTLs viz. *Qsb.pau-2A*, *Qsb.pau-2B*, *Qsb.pau-3B*, and *Qsb.pau-5B*. Thus, there is a need to explore more marker systems to design markers for marker-assisted transfer of other QTLs identified in the present study.

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: <https://www.ncbi.nlm.nih.gov/bioproject/722517>, PRJNA722517.

REFERENCES

- Adhikari, T. B., Gurung, S., Hansen, J. M., Jackson, E. W., and Bonman, J. M. (2012). Association mapping of quantitative trait loci in spring wheat landraces conferring resistance to bacterial leaf streak and spot blotch. *Plant Genome* 5, 1–16. doi: 10.3835/plantgenome2011.12.0032
- Ali, M. S., and Baek, K. H. (2020). Protective roles of cytosolic and plastidial proteasomes on abiotic stress and pathogen invasion. *Plants* 9, 1–17. doi: 10.3390/plants9070832
- Awlchew, Z. T., Singh, R., Kaur, S., Bains, N. S., and Chhuneja, P. (2016). Transfer and mapping of the heat tolerance component traits of *Aegilops speltoides* in tetraploid wheat *Triticum durum*. *Mol. Breed.* 36, 78–92. doi: 10.1007/s11032-016-0499-2
- Ayana, G. T., Ali, S., Sidhu, J. S., Gonzalez Hernandez, J. L., Turnipseed, B., and Sehgal, S. K. (2018). Genome-wide association study for spot blotch resistance in hard winter wheat. *Front. Plant Sci.* 9:926.
- Bainsla, N. K., Phuke, R. M., He, X., Gupta, V., Bishnoi, S. K., Sharma, R. K., et al. (2020). Genome-wide association study for spot blotch resistance in Afghan wheat germplasm. *Plant Pathol.* 69, 1161–1171. doi: 10.1111/ppa.13191
- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Beth Mudgett, M. (2005). New insights to the function of phytopathogenic bacterial type III effectors in plants. *Annu. Rev. Plant Biol.* 56, 509–531. doi: 10.1146/annurev.arplant.56.032604.144218
- Calderón-Villalobos, L. I. A., Nill, C., Marrocco, K., Kretsch, T., and Schwechheimer, C. (2007). The evolutionarily conserved Arabidopsis thaliana F-box protein AtFBP7 is required for efficient translation during temperature stress. *Gene* 392, 106–116. doi: 10.1016/j.gene.2006.11.016
- Chand, R., Pandey, S. P., Singh, H. V., Kumar, S., and Joshi, A. K. (2003). Variability and its probable cause in natural populations of spot blotch pathogen *Bipolaris sorokiniana* of wheat (*T. aestivum* L.) in India / Variabilität und ihre wahrscheinliche Ursache in natürlichen Populationen des Erregers von Blattflecken (*Bipolaris*). *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz / J. Plant Dis. Prot.* 110, 27–35.
- Chisholm, S. T., Coaker, G., Day, B., and Staskawicz, B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124, 803–814. doi: 10.1016/j.cell.2006.02.008

AUTHOR CONTRIBUTIONS

JKp and PC: design and supervision of the study. JKp, SK, AS, and PC: generation of plant genetic material. JKh, JKp, RB, HK, and JS: screening of material. JKh and GD: data analysis. JKh, JKp, RB, PS, SK, AS, and PC: management of trial. JKh, JKp, GD, and PC: draft of manuscript. All authors contributed for compilation of the article.

FUNDING

This work was supported by Science and Engineering Research Board (SERB), Department of Science & Technology (DST), Government of India with grant no. EMR/2016/002474.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.650400/full#supplementary-material>

- Chowdhury, A. K., Singh, G., Tyagi, B. S., Ojha, A., Dhar, T., and Bhattacharya, P. M. (2013). Spot blotch disease of wheat – a new thrust area for sustaining productivity. *J. Wheat Res.* 5, 1–11.
- Dhillon, G. S., Kaur, S., Das, N., Singh, R., Poland, J., Kaur, J., et al. (2020). QTL mapping for stripe rust and powdery mildew resistance in *Triticum durum* - *Aegilops speltoides* backcross introgression lines. *Plant Genet. Resour. Characterisation Util.* 18, 211–221. doi: 10.1017/S1479262120000222
- Dubey, N., and Singh, K. (2018). “Role of NBS-LRR proteins in plant defense,” in *Molecular Aspects of Plant-Pathogen Interaction*, eds A. Singh and I. K. Singh (Singapore: Springer), 115–138. doi: 10.1007/978-981-10-7371-7_5
- Duveiller, E., and Sharma, R. C. (2012). “Wheat resistance to spot blotch or foliar blight,” in *Disease Resistance in Wheat*, ed. I. Sharma (Wallingford: CAB), 120–135. doi: 10.1079/9781845938185.0120
- Fu, J., and Wang, S. (2011). Insights into auxin signaling in plant-pathogen interactions. *Front. Plant Sci.* 2:74. doi: 10.3389/fpls.2011.00074
- Glaubit, J. C., Castevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS One* 9:e90346. doi: 10.1371/journal.pone.0090346
- Gunupuru, L. R., Arunachalam, C., Malla, K. B., Kahla, A., Perochon, A., Jia, J., et al. (2018). A wheat cytochrome P450 enhances both resistance to deoxynivalenol and grain yield. *PLoS One* 13:e0204992. doi: 10.1371/journal.pone.0204992
- Gupta, P. K., Chand, R., Vasistha, N. K., Pandey, S. P., Kumar, U., Mishra, V. K., et al. (2018). Spot blotch disease of wheat: the current status of research on genetics and breeding. *Plant Pathol.* 67, 508–531. doi: 10.1111/ppa.12781
- Gurung, S., Mamidi, S., Bonman, J. M., Xiong, M., Brown-Guedira, G., and Adhikari, T. B. (2014). Genome-Wide Association Study Reveals Novel Quantitative Trait Loci Associated with Resistance to Multiple Leaf Spot Diseases of Spring Wheat. *PLoS One* 9:e108179. doi: 10.1371/journal.pone.0108179
- Jamil, M., Ali, A., Gul, A., Ghafoor, A., Ibrahim, A. M. H., and Mujeib-Kazi, A. (2018). Genome-wide association studies for spot blotch (*Cochliobolus sativus*) resistance in bread wheat using genotyping-by-sequencing. *Phytopathology* 108, 1307–1314. doi: 10.1094/phyto-02-18-0047-r
- Joshi, A. K., Kumar, S., Chand, R., and Ortiz-Ferrera, G. (2004). Inheritance of resistance to spot blotch caused by *Bipolaris sorokiniana* in spring wheat. *Plant Breed.* 123, 213–219. doi: 10.1111/j.1439-0523.2004.00954.x
- Joshi, A. K., Kumari, M., Singh, V. P., Reddy, C. M., Kumar, S., Rane, J., et al. (2006). Stay green trait: variation, inheritance and its association with spot blotch resistance in spring wheat (*Triticum aestivum* L.). *Euphytica* 153, 59–71. doi: 10.1007/s10681-006-9235-z

- Joshi, A. K., Mishra, B., Chatrath, R., Ortiz Ferrara, G., and Singh, R. P. (2007). Wheat improvement in India: present status, emerging challenges and future prospects. *Euphytica* 157, 431–446. doi: 10.1007/s10681-007-9385-7
- Kassambara, A. (2020). *ggpubr: 'ggplot2' Based Publication Ready Plots*. 155. Available online at: <https://rpkgs.datanovia.com/ggpubr/> (accessed August 23, 2020).
- Kassambara, A., and Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. 1–84. Available online at: <https://cran.r-project.org/package=factoextra> (accessed May 05, 2020).
- Kaur, S., Jindal, S., Kaur, M., and Chhuneja, P. (2018). “Utilization of wild species for wheat improvement using genomic approaches,” in *Biotechnologies of Crop Improvement, Volume 3: Genomic Approaches*, eds S. S. Gosal and S. H. Wani (Cham: Springer International Publishing), 105–150. doi: 10.1007/978-3-319-94746-4_6
- Khan, H., and Chowdhury, S. (2011). Identification of resistance source in wheat germplasm against spot blotch disease caused by *Bipolaris sorokiniana*. *Arch. Phytopathol. Plant Prot.* 44, 840–844. doi: 10.1080/03235400903345273
- Kim, D. H., Yamaguchi, S., Lim, S., Oh, E., Park, J., Hanada, A., et al. (2008). SOMNUS, a CCCH-type zinc finger protein in Arabidopsis, negatively regulates light-dependent seed germination downstream of PIL5. *Plant Cell* 20, 1260–1277. doi: 10.1105/tpc.108.058859
- Kim, H. S., and Delaney, T. P. (2002). Arabidopsis SON1 is an F-box protein that regulates a novel induced defense response independent of both salicylic acid and systemic acquired resistance. *Plant Cell* 14, 1469–1482. doi: 10.1105/tpc.001867
- Kumar, S., Röder, M. S., Tripathi, S. B., Kumar, S., Chand, R., Joshi, A. K., et al. (2015). Mendelization and fine mapping of a bread wheat spot blotch disease resistance QTL. *Mol. Breed.* 35, 218–223. doi: 10.1007/s11032-015-0411-5
- Kumar, U., Joshi, A. K., Kumar, S., Chand, R., and Röder, M. S. (2008). Mapping of resistance to spot blotch disease caused by *Bipolaris sorokiniana* in spring wheat. *Theor. Appl. Genet.* 118, 783–792. doi: 10.1007/s00122-008-0938-5
- Kumar, U., Joshi, A. K., Kumar, S., Chand, R., and Röder, M. S. (2010). Quantitative trait loci for resistance to spot blotch caused by *Bipolaris sorokiniana* in wheat (*T. aestivum* L.) lines ‘Ning 8201’ and ‘Chirya 3’. *Mol. Breed.* 26, 477–491. doi: 10.1007/s11032-009-9388-2
- Ladejobi, O., Mackay, I. J., Poland, J., Proud, S., Hibberd, J. M., and Bentley, A. R. (2019). Reference genome anchoring of high-density markers for association mapping and genomic prediction in European winter wheat. *Front. Plant Sci.* 10:1278. doi: 10.3389/fpls.2019.01278
- Latwal, C., Deepshikha, Kumari, B., Singh, P. K., and Jaiswal, J. P. (2016). Characterization of bread wheat germplasm for spot blotch resistance and its association with yield and yield related traits. *J. Wheat Res.* 8, 31–37.
- Lê, S., Josse, J., and Husson, F. (2008). Facto Mine R: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01
- Lee, H. A., and Yeom, S. I. (2015). Plant NB-LRR proteins: tightly regulated sensors in a complex manner. *Brief. Funct. Genomics* 14, 233–242. doi: 10.1093/bfpg/elt012
- Levy, M., Wang, Q., Kaspi, R., Parrella, M. P., and Abel, S. (2005). Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *Plant J.* 43, 79–96. doi: 10.1111/j.1365-313x.2005.02435.x
- Lillemo, M., Joshi, A. K., Prasad, R., Chand, R., and Singh, R. P. (2013). QTL for spot blotch resistance in bread wheat line Saar co-locate to the biotrophic disease resistance loci Lr34 and Lr46. *Theor. Appl. Genet.* 126, 711–719. doi: 10.1007/s00122-012-2012-6
- Lin, P. C., Pomeranz, M. C., Jikumaru, Y., Kang, S. G., Hah, C., Fujioka, S., et al. (2011). The Arabidopsis tandem zinc finger protein AtTZF1 affects ABA- and GA-mediated growth, stress and gene expression responses. *Plant J.* 65, 253–268. doi: 10.1111/j.1365-313X.2010.04419.x
- Llorente, F., Muskett, P., Sánchez-Vallet, A., López, G., Ramos, B., Sánchez-Rodríguez, C., et al. (2008). Repression of the auxin response pathway increases arabidopsis susceptibility to necrotrophic fungi. *Mol. Plant* 1, 496–509. doi: 10.1093/mp/ssn025
- Lu, D., Lin, W., Gao, X., Wu, S., Cheng, C., Avila, J., et al. (2011). Direct ubiquitination of pattern recognition receptor FLS2 attenuates plant innate immunity. *Science* 332, 1439–1442. doi: 10.1126/science.1204903
- Lu, P., Liang, Y., Li, D., Wang, Z., Li, W., Wang, G., et al. (2016). Fine genetic mapping of spot blotch resistance gene Sb3 in wheat (*Triticum aestivum*). *Theor. Appl. Genet.* 129, 577–589. doi: 10.1007/s00122-015-2649-z
- Ma, L., and Li, G. (2018). Far1-Related Sequence (FRS) and frs-related factor (FRF) family proteins in arabidopsis growth and development. *Front. Plant Sci.* 9:692. doi: 10.3389/fpls.2018.00692
- Ma, L., Tian, T., Lin, R., Deng, X.-W., Wang, H., and Li, G. (2016). Arabidopsis FHY3 and FAR1 regulate light-induced myo -inositol biosynthesis and oxidative stress responses by transcriptional activation of MIPS1. *Mol. Plant* 9, 541–557. doi: 10.1016/j.molp.2015.12.013
- Maldonado-Bonilla, L. D., Eschen-Lippold, L., Gago-Zachert, S., Tabassum, N., Bauer, N., Scheel, D., et al. (2013). The arabidopsis tandem zinc finger 9 protein binds rna and mediates pathogen-associated molecular pattern-triggered immune responses. *Plant Cell Physiol.* 55, 412–425. doi: 10.1093/pcp/pct175
- Marino, D., Peeters, N., and Rivas, S. (2012). Ubiquitination during plant immune signaling. *Plant Physiol.* 160, 15–27. doi: 10.1104/pp.112.199281
- Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL Ici mapping: 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
- Montalbini, P. (1991). Effect of rust infection on levels of uricase, allantoinase and ureides in susceptible and hypersensitive bean leaves. *Physiol. Mol. Plant Pathol.* 39, 173–188. doi: 10.1016/0885-5765(91)90002-Y
- Neupane, A., Sharma, R., Duveiller, E., and Shrestha, S. (2010). Sources of Cochliobolus sativus inoculum causing spot blotch under warm wheat growing conditions in South Asia. *Cereal Res. Commun.* 38, 541–549. doi: 10.1556/CRC.38.2010.4.11
- R Core Team. (2019). *A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reuber, T. L., and Ausubel, F. M. (1996). Isolation of Arabidopsis genes that differentiate between resistance responses mediated by the RPS2 and RPM1 disease resistance genes. *Plant Cell* 8, 241–249. doi: 10.1105/tpc.8.2.241
- Rosyara, U. R., Khadka, K., Subedi, S., Sharma, R. C., and Duveiller, E. (2008). Heritability of stay green traits and association with spot blotch resistance in three spring wheat populations. *J. Genet. Breed.* 62, 75–82.
- Saari, E. E. (1998). “Leaf blight diseases and associated soilborne fungal pathogens of wheat in South and Southeast Asia,” in *Helminthosporium Blights Wheat Spot Blotch Tan Spot*, eds E. Duveiller, H. J. Dubin, J. Reeves, and A. McNab 37–51.
- Sadanandom, A., Bailey, M., Ewan, R., Lee, J., and Nelis, S. (2012). The ubiquitin-proteasome system: central modifier of plant signalling. *New Phytol.* 196, 13–28. doi: 10.1111/j.1469-8137.2012.04266.x
- Saghai-Marouf, M. A., Soliman, K. M., Jorgensen, R. A., and Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 81, 8014–8018. doi: 10.1073/pnas.81.24.8014
- Shaner, G., and Finney, R. E. (1977). The effect of nitrogen fertilization on the expression of slow-mildewing resistance in knox wheat. *Phytopathology* 77, 1051–1056. doi: 10.1094/phyto-67-1051
- Singh, D. P., and Kumar, P. (2008). Role of spot blotch (*Bipolaris sorokiniana*) in deteriorating seed quality, its management in different wheat genotypes using fungicidal seed treatment. *Indian Phytopathol.* 61, 49–54.
- Singh, P. J., and Dhaliwal, H. S. (1993). Resistance to foliar blight of wheat in Agilops and wild Triticum species. *Indian Phytopathol.* 46, 246–248.
- Singh, P. K., He, X., Sansaloni, C. P., Juliana, P., Dreisigacker, S., Duveiller, E., et al. (2018). Resistance to spot blotch in two mapping populations of common wheat is controlled by multiple QTL of minor effects. *Int. J. Mol. Sci.* 19:4054. doi: 10.3390/ijms19124054
- Smurova, S. G., and Mikhailova, L. A. (2007). Sources of resistance to wheat spot blotch. *Russ. Agric. Sci.* 33, 378–380. doi: 10.3103/s1068367407060092
- Sun, J., Jiang, H., Xu, Y., Li, H., Wu, X., Xie, Q., et al. (2007). The CCCH-type zinc finger proteins AtSZF1 and AtSZF2 regulate salt stress responses in arabidopsis. *Plant Cell Physiol.* 48, 1148–1158. doi: 10.1093/pcp/pcm088
- Thomas, E. L., and van der Hoorn, R. A. L. (2018). Ten prominent host proteases in plant-pathogen interactions. *Int. J. Mol. Sci.* 19, 639–653. doi: 10.3390/ijms19020639
- Tomar, V., Singh, D., Dhillon, G. S., Singh, R. P., Poland, J., Joshi, A. K., et al. (2021). New QTLs for spot blotch disease resistance in wheat (*Triticum*

- aestivum* L.) using genome-wide association mapping. *Front. Genet.* 11:613217. doi: 10.3389/fgene.2020.613217
- Trujillo, M., and Shirasu, K. (2010). Ubiquitination in plant immunity. *Curr. Opin. Plant Biol.* 13, 402–408. doi: 10.1016/j.pbi.2010.04.002
- Üstün, S., Sheikh, A., Gimenez-Ibanez, S., Jones, A., Ntoukakis, V., and Börnke, F. (2016). The proteasome acts as a hub for plant immunity and is targeted by *Pseudomonas* type III effectors. *Plant Physiol.* 172, 1941–1958. doi: 10.1104/pp.16.00808
- Wickham, H. ed. (2016). *ggplot2: elegant graphics for data analysis*. 2nd ed New York: Springer-Verlag Available online at: <https://ggplot2.tidyverse.org> (accessed January 05, 2021).
- Woo, H. R., Chung, K. M., Park, J. H., Oh, S. A., Ahn, T., Hong, S. H., et al. (2001). ORE9, an F-box protein that regulates leaf senescence in Arabidopsis. *Plant Cell* 13, 1779–1790. doi: 10.1105/tpc.13.8.1779
- Yadav, B., Singh, R., and Kumar, A. (2015). Management of spot blotch of wheat using Fungicides. Bio-agents and Botanicals. *Afr. J. Agric. Res.* 10, 2494–2500. doi: 10.5897/AJAR2013.8196
- Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x
- Zhang, P., Guo, G., Wu, Q., Chen, Y., Xie, J., Lu, P., et al. (2020). Identification and fine mapping of spot blotch (*Bipolaris sorokiniana*) resistance gene Sb4 in wheat. *Theor. Appl. Genet.* 133, 2451–2459. doi: 10.1007/s00122-020-03610-3
- Zhuang, Y., Gala, A., and Yen, Y. (2013). Identification of functional genic components of major Fusarium head blight resistance quantitative trait loci in wheat cultivar Sumai 3. *Mol. Plant Microbe Interact.* 26, 442–450. doi: 10.1094/mpmi-10-12-0235-r

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 Kaur, Kaur, Dhillon, Kaur, Singh, Bala, Srivastava, Kaur, Sharma and Chhuneja. 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 Association Study and Genomic Prediction for Soybean Cyst Nematode Resistance in USDA Common Bean (*Phaseolus vulgaris*) Core Collection

Ainong Shi^{1*}, Paul Gepts², Qijian Song³, Haizheng Xiong¹, Thomas E. Michaels⁴ and Senyu Chen^{5*}

¹ Department of Horticulture, PTSC316, University of Arkansas, Fayetteville, AR, United States, ² Department of Plant Sciences, University of California, Davis, CA, United States, ³ United States Department of Agriculture, Agricultural Research Service, Beltsville Agricultural Research Center, Beltsville, MD, United States, ⁴ Department of Horticultural Science, University of Minnesota, St. Paul, MN, United States, ⁵ Southern Research and Outreach Center, University of Minnesota, Waseca, MN, United States

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Andrés J. Cortés,
Colombian Corporation for Agricultural
Research (AGROSAVIA), Colombia
Subas Malla,
Texas A&M University, United States
Dapeng Zhang,
Beltsville Agricultural Research Center,
United States

*Correspondence:

Senyu Chen
chenx099@umn.edu
Ainong Shi
ashi@uark.edu

Specialty section:

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

Received: 30 October 2020

Accepted: 14 May 2021

Published: 07 June 2021

Citation:

Shi A, Gepts P, Song Q, Xiong H,
Michaels TE and Chen S (2021)
Genome-Wide Association Study and
Genomic Prediction for Soybean Cyst
Nematode Resistance in USDA
Common Bean (*Phaseolus vulgaris*)
Core Collection.
Front. Plant Sci. 12:624156.
doi: 10.3389/fpls.2021.624156

Soybean cyst nematode (SCN, *Heterodera glycines*) has become the major yield-limiting biological factor in soybean production. Common bean is also a good host of SCN, and its production is challenged by this emerging pest in many regions such as the upper Midwest USA. The use of host genetic resistance has been the most effective and environmentally friendly method to manage SCN. The objectives of this study were to evaluate the SCN resistance in the USDA common bean core collection and conduct a genome-wide association study (GWAS) of single nucleotide polymorphism (SNP) markers with SCN resistance. A total of 315 accessions of the USDA common bean core collection were evaluated for resistance to SCN HG Type 0 (race 6). The common bean core set was genotyped with the BARCBean6K_3 Infinium BeadChips, consisting of 4,654 SNPs. Results showed that 15 accessions were resistant to SCN with a Female Index (FI) at 4.8 to 9.4, and 62 accessions were moderately resistant ($10 < FI < 30$) to HG Type 0. The association study showed that 11 SNP markers, located on chromosomes Pv04, 07, 09, and 11, were strongly associated with resistance to HG Type 0. GWAS was also conducted for resistance to HG Type 2.5.7 and HG Type 1.2.3.5.6.7 based on the public dataset ($N = 276$), consisting of a diverse set of common bean accessions genotyped with the BARCBean6K_3 chip. Six SNPs associated with HG Type 2.5.7 resistance on Pv 01, 02, 03, and 07, and 12 SNPs with HG Type 1.2.3.5.6.7 resistance on Pv 01, 03, 06, 07, 09, 10, and 11 were detected. The accuracy of genomic prediction (GP) was 0.36 to 0.49 for resistance to the three SCN HG types, indicating that genomic selection (GS) of SCN resistance is feasible. This study provides basic information for developing SCN-resistant common bean cultivars, using the USDA core germ plasm accessions. The SNP markers can be used in molecular breeding in common beans through marker-assisted selection (MAS) and GS.

Keywords: common bean, *Phaseolus vulgaris*, soybean cyst nematode, *Heterodera glycines*, genomic prediction, genome wide association study, genomic selection, single nucleotide polymorphism

INTRODUCTION

Common bean (*Phaseolus vulgaris* L.) is the most important edible grain legume crop worldwide, with crop value equal to the combined value of all other food legumes such as peas and chickpeas (Jain et al., 2016). The most common bean is harvested as seed grain called “dry bean,” but it is also grown as a green vegetable (called “green bean” or “snap bean”) in many parts of the world. Common bean has high nutritional value and is one of the most important sources of protein for billions of people in the world. In recent years, about 2 million acres were planted, and approximately 1.3 million metric tons of common beans valued at US\$2 billion were produced annually in the United States (US) (USDA NASS, 2020).

The production of dry beans in the US may be challenged by an emerging, invasive pest, the soybean cyst nematode (SCN), *Heterodera glycines* Ichinohe (Tylenchida: Heteroderidae). The SCN is the most serious pathogen of soybean [*Glycine max* (L.) Merr.] in the US and suppresses a yield more than any other pathogen (Koenning and Wrather, 2010; Allen et al., 2017). The SCN reduces a yield by feeding on plant nutrients, retarding root growth, reducing water and nutrient uptake and transportation from roots to shoots, and inhibiting rhizobium nodulation. Yield losses can exceed 40% (Koenning and Wrather, 2010), depending on many factors such as SCN population density, soil texture and fertility, rainfall, and the presence of susceptible soybean genotypes (Duan et al., 2009). The SCN has been widely spread in the US, especially in the North Central region that produces most soybeans (Tylka and Maret, 2017). Unfortunately, the top four common bean-growing states, North Dakota, Michigan, Nebraska, and Minnesota, which produce approximately 70% of the common bean in the US, are also in the North Central region. The SCN has been reported in the common bean fields of those states (Poromarto et al., 2010; Yan et al., 2017). SCN infection can cause severe yield loss without any aboveground symptoms in common beans (Poromarto et al., 2010, 2012) and becomes a serious threat to common bean production.

The use of host resistance has been highly successful in SCN management for soybeans. Numerous commercial SCN-resistant soybean cultivars are available and are planted in most soybean fields in the US. Similarly, the use of host resistance in common bean cultivars will also be crucial to SCN management in dry bean production. Growing common bean cultivars resistant to SCN infection will not only reduce common bean yield loss but also relieve SCN pressure for soybean production if common beans and soybean are rotated with wheat (*Triticum aestivum* L.). Recently, Osorno et al. (2020) has released the first pinto bean cultivar “ND Falcon,” a new pinto bean with combined resistance to SCN and rust. Screening more common bean germplasm for SCN resistance, using different HG Types (races) will provide breeders to use germplasm as parents to develop and release new superior common bean cultivars with broad and more stable resistance.

Limited research has demonstrated that some common bean germplasm and cultivars are resistant to SCN. Smith and Young (2003) conducted a greenhouse study to evaluate 20 common bean lines for SCN resistance and found a few lines resistant to

SCN, and some Mesoamerican genotypes were more resistant than Andean genotypes. Poromarto et al. (2012), in North Dakota, evaluated 416 accessions (germplasm lines) in the USDA core collection of *P. vulgaris* and found 23% of the lines had low nematode reproduction and were considered highly resistant to SCN HG Type 0 (Jain et al., 2016, 2019). Wen et al. (2019), in Illinois, evaluated 363 accessions of the same core collection and found 19 accessions (5%) were highly resistant to SCN HG Type 2.5.7, and 160 (44%) resistant to HG Type 1.2.3.5.6.7, with FI < 10.

Jain et al. (2016) analyzed the transcriptome sequences of the SCN-resistant line PI533561 vs. SCN-susceptible *P. vulgaris* line GTS-900 and demonstrated that genes-encoding nucleotide-binding site leucine-rich repeat resistance (NLR) proteins, WRKY transcription factors, pathogenesis-related (PR) proteins, and heat shock proteins involved in diverse biological processes were differentially expressed between SCN-resistant and susceptible genotypes. Recently, two reports on SCN-resistant quantitative trait loci (QTLs) in common beans were published. Wen et al. (2019) conducted a genome-wide association study (GWAS) based on the dataset of 363 USDA common bean core accessions phenotyped against SCN HG types 2.5.7 and 1.2.3.5.6.7 and genotyped, using 84,416 single nucleotide polymorphisms (SNPs) obtained from genotyping by sequencing (GBS) and reported that there were five SNPs on chromosome Pv01 and one on Pv09 associated with resistance to HG Type 2.5.7. They also reported a gene cluster orthologous to the three genes at the SCN-resistant *rhg1* locus in soybeans. In addition, an SNP was found on Pv09, associated with resistance to HG Type 1.2.3.5.6.7. Jain et al. (2019) conducted GWAS in 317 accessions of USDA common bean core collection, challenged with SCN HG Type 0, and found 14 significant SNP markers on Pv04, 05, 06, 07, 08, 10, and 11 in the Middle American subpopulation and 23 SNP markers on Pv01, 02, 07, 08, 09, and 11 for the Andean subpopulation. Besides, Jain et al. (2019) reported several candidate genes on Pv01 and Pv08, which had high similarity to the three genes of *rhg1* of soybean for SCN resistance. Based on previous reports and the study, the SCN resistance in the common bean is polygenic traits with multiple genes or alleles.

Plant molecular breeding has been the foundation for crop improvement into the twenty first century and has become part of the breeding programs to expedite advances and genetic gains in many crops (Moose and Mumm, 2008). Marker-assisted selection (MAS) has been successfully used in the selection of specific major genes/alleles in plant breeding (Collard et al., 2005; Collard and Mackill, 2008; Xu and Crouch, 2008). More recently, predictive breeding via GS has become an essential tool in crop improvement. GS refers to selecting the performance of individuals within a population based on genomic-estimated breeding values (GEBV) (Hayes et al., 2009; Desta and Ortiz, 2014). The decreasing cost of DNA sequencing renders GS affordable and powerful by providing high-density markers across the genome (Lin et al., 2014). GS is more efficient than the traditional MAS when dealing with small-effect QTL (Bernardo and Yu, 2007; Heffner et al., 2009, 2011; Cortés et al., 2020). So far, genomic prediction (GP) as a GS parameter has been investigated

in a dozen of crops such as maize (*Zea mays* L.), rice (*Oryza sativa* L.), soybean, and wheat (Bernardo and Yu, 2007; Heffner et al., 2009, 2011; Albrecht et al., 2011; Jarquin et al., 2014, 2016; Onogi et al., 2016; Xavier et al., 2016; Shikha et al., 2017; Zhang et al., 2017; Qin et al., 2019) for various agronomic traits, and abiotic and biotic stress traits. Genomic breeding value estimation in GP is the key step in GS. Several approaches have been proposed for GEBV, such as BLUP methods (gBLUP, RR-BLUP, cBLUP, and sBLUP) and Bayesian methods (BayesA and BayesB). All articles discussed the selection prediction accuracy (PA), estimated using the Pearson's correlation coefficient (r) between the GEBV and observed values for each trait in validation sets (testing sets), using several models. In recent years, GP has also been reported in common beans to predict agronomic traits under different environmental stresses (Keller et al., 2020) and SCN resistance (Wen et al., 2019).

Currently, SNP technology is the molecular-marker platform of choice in genome-wide mapping, association studies, diversity analysis, and tagging of important genes in plant genomics and breeding. SNPs are abundant in the genome, cost-effective, and amenable to high throughput analysis (Collard and Mackill, 2008; Xu and Crouch, 2008). Therefore, the identification of SNP markers will provide breeders with powerful tools to assist in selecting biotic and abiotic stress resistance/tolerance and expedite the development of elite cultivars with stress resistance/tolerance in common bean breeding programs. SNPs have been reported and used in common beans (Cortés et al., 2011; Blair et al., 2013). Gene-based SNP markers were developed in common beans (Galeano et al., 2012). SNP genetic maps for common beans have been constructed, using the 6K SNP BeadChips (Song et al., 2015) and were used to anchor the scaffold of the common bean whole-genome sequence reference assembly for the Andean landrace G19833 (Schmutz et al., 2014). In common beans, the BARCBean6K_3 Infinium BeadChip has been used for QTL and association mapping to identify genes/QTL controlling different traits (Hagerty et al., 2015, 2016; Hoyos-Villegas et al., 2016, 2017; Moghaddam et al., 2016; Castro et al., 2017; Hurtado-Gonzales et al., 2017; Valentini et al., 2017). Recently, several versions of *P. vulgaris* (common bean) genome assemblies were released. They include the aforementioned Andean genome (Schmutz et al., 2014; https://phytozome-next.jgi.doe.gov/info/Pvulgaris_v2_1) and four Middle American genomes: (1) race Mesoamerica: cultivar OAC Rex (https://www.ncbi.nlm.nih.gov/genome/380?genome_assembly_id=1500596) and breeding line BAT93 (https://www.ncbi.nlm.nih.gov/genome/380?genome_assembly_id=262776; Vlasova et al., 2016; Rendón-Anaya et al., 2017); (2) race Durango: cultivar Pinto UI111 (https://phytozome-next.jgi.doe.gov/info/PvulgarisUI111_v1_1), and cultivar Labor Ovalle of race Guatemala (https://phytozome-next.jgi.doe.gov/info/PvulgarisLaborOvalle_v1_1). The genome assembly of G19833 has been used as a reference to map SNPs of the BARCBean6K_3 Infinium BeadChip to the 11 chromosomes in common beans (Song et al., 2015).

With the decreased genotyping cost and improved statistical methods, GWAS and GS offer new approaches for genetic improvement of complex traits in crop species. GWAS, based on a population of unrelated lines and high-density markers,

has been used to identify candidate genes for a broad range of complex traits in different crops (Huang et al., 2010; Li et al., 2013; Morris et al., 2013; Yano et al., 2016). GWAS is relatively new for common beans, but it has been reported to be an effective approach to identify SNP markers associated with SCN resistance (Jain et al., 2019; Wen et al., 2019). However, MAS has been successfully coupled with backcrossing schemes for transferring several traits, among which anthracnose resistance and seed mineral accumulation traits (even from the wild) in common beans (Garzón et al., 2008; Blair and Izquierdo, 2012). Therefore, research is needed to identify SNP markers associated with SCN resistance and to use these SNP markers in molecular breeding to enhance common bean improvement.

We initiated a project in 2016 to study the SCN resistance in common beans, using SCN HG Type 0. So far, two studies for SCN resistance QTLs in the USDA common bean core collection have been reported (Jain et al., 2019; Wen et al., 2019). Wen et al. (2019) conducted GWAS in 363 accessions of USDA common bean core collection phenotyped against SCN HG Types 2.5.7 and 1.2.3.5.6.7 and genotyped, using GBS. The common bean core sets were also genotyped BARCBean6K_3 Infinium BeadChip SNPs, and the SNP data are available (Song et al., 2015; Gepts et al., 2019; Kuzay et al., 2020). The BARCBean6K_3 Infinium BeadChip could provide additional SNPs for a breeding program. Therefore, we conduct GWAS and GP analysis for resistance to HG Type 2.5.7 and HG Type 1.2.3.5.6.7, using the phenotypic data from Wen et al. (2019) and genotypic data of the BARCBean6K chip SNPs in this report. Although Jain et al. (2019) conducted GWAS in 317 accessions of USDA common bean core collection with SCN HG Type 0, only 86 accessions with FI < 10 were published in the article; hence, their data are not included in this study for further analysis. The overall goal of the research was to develop technology to effectively manage the SCN in common bean productions. Specifically, the objectives of this study were to evaluate the SCN resistance in the USDA common bean core collection, to conduct GWAS, and to identify SNP markers associated with SCN resistance. The approach was to first conduct GWAS to identify associated SNP markers and then use the associated SNP markers to do GS. This is an approach combining MAS and GS through GEBVs, using associated SNP markers (Spindel et al., 2016; Zhang J. P. et al., 2016; Qin et al., 2019; Ravelombola et al., 2019, 2020, 2021; Ali et al., 2020). The information presented in this report is a new contribution to the understanding of SCN resistance in common beans beyond the previous studies (Jain et al., 2019; Wen et al., 2019).

MATERIALS AND METHODS

Plant Materials

About 315 common bean germplasm accessions, a core set of common beans, described at USDA Germplasm Resources Information Network (GRIN), were used in this study. This common bean core set has been widely used for genetic diversity analysis (Kwak and Gepts, 2009; McClean et al., 2012; Campa et al., 2018; Gepts et al., 2019; Kuzay et al., 2020). The core set was mainly from two gene pools, i.e., the Andean and Mesoamerican pools (Gepts et al., 1986, 2019; Koenig and Gepts,

1989; Koinange and Gepts, 1992; Beebe et al., 1997, 2000; Kwak and Gepts, 2009; Bitocchi et al., 2012, 2013; McClean et al., 2012; Schmutz et al., 2014; Campa et al., 2018; Kuzay et al., 2020), and can form three clusters and seven groups (Kuzay et al., 2020). The 315 accessions in this study were originally collected from 11 countries, including Mexico (163 accessions), Colombia (35), Guatemala (30), Peru (17), Costa Rica (17), Ecuador (16), El Salvador (13), Nicaragua (13), Honduras (9), Bolivia (1), and United States (1) (**Supplementary Table 1**). They represented 241 accessions from Middle American gene pools, 67 from the Andean pool, and seven from an admixture (**Supplementary Table 1**).

In addition, the seven soybean SCN HG Type indicator (differential) lines PI 548402 (Peking), PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, and PI 548316 (Niblack et al., 2002), and four SCN race differential lines PI 548402 (Peking), PI 548982 (Pickett 71), or PI 548988 (Pickett), PI 88788, and PI 90763 (Riggs and Schmitt, 1988), with the susceptible Williams 82 (PI 518671) as control were included to confirm the virulence phenotype of the SCN population (**Supplementary Table 2**). Based on the reactions of the differential lines to the SCN population, the population was HG Type 0 and race 6 similar to race 3.

Soybean Cyst Nematode Resistance Phenotyping

The 315 common bean accessions were tested for their resistance to SCN HG Type 0 (race 6). HG Type 0 is avirulent to most current commercial SCN-resistant soybean cultivars, and if there is any SCN resistance in common beans, it is likely resistant to HG Type 0 based on the knowledge of SCN resistance in soybeans. Consequently, we started screening, using the HG Type 0, to identify more SCN-resistant common bean lines and genes/alleles.

The SCN population was collected from a field in Swift County, Minnesota, USA, in 2007. Since it was collected from the field, the population had been maintained in the greenhouse on susceptible soybean cultivars or stored in a freezer at -20°C . Prior to the experiment, the nematode population was cultured on susceptible soybean “sturdy” for about 45 days. Newly formed females and cysts were washed with a vigorously applied water stream through an 850- μm -aperture sieve onto a 250- μm -aperture sieve and extracted by centrifugation in a 63% (w/v) sucrose solution. Eggs were released from the cysts by crushing the cysts on a 150- μm -aperture sieve with a rubber stopper mounted on a motor (Faghihi and Ferris, 2000). The eggs were separated from debris by centrifugation in a 35% (w/v) sucrose solution for 5 min at 1,500 g, and an egg suspension of 800 eggs/ml was made. The reproduction of the SCN population on the soybean or bean lines was assayed by growing the bean in cone-tainers (4-cm diameter and 13.5-cm high) in a growth room (**Supplementary Figure 1**).

The experimental design was a randomized complete block design (RCBD) with three replicates. Each replicate included two common bean plants in two separate cone-tainers per common bean accession. Control soybean Williams 82 in each

replicate included five plants in five separate cone-tainers. All three replicates of the 315 common bean accessions, with a total of 1,890 cone-tainers, plus the Williams 82, were set up within 2 days of December 14 to 15, 2016, in the growth room, and they were arranged in three blocks (**Supplementary Figure 1**). The cone-tainers were filled with autoclaved soil (80% sands + 20% field clay loam soil) to half to which 2,000 eggs in 2.5 ml of water were added. Additional soil was placed in the cone-tainer to approximately 2 cm from the top. Another inoculum of 2,000 eggs in 2.5 ml of water was added to the soil surface, and one common bean or soybean seed was sowed in each cone-tainer. The seed was covered with additional soil to about 1-cm depth. The cone-tainers were placed on a rack and maintained in the growth room with the temperature set at 28°C and daily artificial lights for 16 h. Water was applied with a sprinkler irrigation system to maintain adequate soil moisture (**Supplementary Figure 1**). The environments, including soil temperature, moisture, and lights, were controlled relatively even over time and across the benches in the growth room. After 35 days in the growth room, the plants were cut to about 1 cm above the soil surface, and all of the cone-tainers were moved to a cool room (4°C) to stop SCN development. The samples were stored in the cool room until processed.

Cysts (females) were extracted from the roots and soil according to established procedures after 35 days. Briefly, the soil and plant roots were removed from the cone-tainer and placed in a beaker, and water was added. Any cysts on the wall of the cone-tainer were washed off. Plant roots were removed and females washed off on an 850- μm -aperture sieve, nested on a 250- μm -aperture sieve. In addition, the cysts in the soil were extracted by pouring soil suspension on the sieves. After rinsing the materials on the 850- μm -aperture sieve, the cysts with debris on the 250- μm -aperture sieve were collected. The cysts were separated from the debris by flotation centrifugation in sucrose solution (63%) and counted under a dissecting microscope.

A FI for each common bean plant was determined by comparing SCN female number of the plant with the average female number on the five plants of Williams 82: $\text{FI} = \text{female number on a given plant} \times 100 / \text{mean number of females on Williams 82}$, where we defined FI on Williams 82 as 100. The average FI of the two plants in each block was considered as one replicate, and three replicates were included.

So far, two studies for SCN resistance in the USDA common bean core collection have been reported (Jain et al., 2019; Wen et al., 2019). Wen et al. (2019) conducted GWAS in 363 accessions of USDA common bean core collection phenotyped against SCN HG Types 2.5.7 and 1.2.3.5.6.7. Among the 363 accessions reported in Wen et al. (2019), 276 accessions were further analyzed for GWAS and GP in this report based on available SNP data. Therefore, we also include their SCN FI data in this study for comparative data analysis. Although Jain et al. (2019) conducted GWAS in 317 accessions of USDA common bean core collection with SCN HG Type 0, only 86 accessions with $\text{FI} < 10$ were published in the article; hence, their data are not included in this study for further analysis.

Phenotypic Data Analysis

The SCN resistance phenotypic data FI of SCN HG Type 0 (race 6) among the 315 common bean accessions were analyzed, using the ANOVA, with the general linear models (GLM) procedure of JMP Genomics 7 (SAS Institute, Cary, NC). For comparisons among individual accessions in JMP, the “LSMeans Student’s *t*” was used to perform multiple comparisons at $\alpha = 0.05$. The mean, range, SD, SE, and coefficient of variation (CV) were estimated for FI, using “Tabulate.” Person’s correlation coefficients (*r*) were calculated, using “Multivariate Methods.” The distribution of FI was drawn, using “Distribution” in JMP Genomics 7. The average of FI to SCN HG Type 0 (race 6) for each soybean accession from ANOVA was used as the phenotypic data for GWAS.

The broad-sense heritability (*H*) was estimated, using the following formula (Holland, 2003).

$$H = \sigma_g^2 / [\sigma_g^2 + (\sigma_e^2/r)]$$

with σ_g^2 being the total genetic variance, σ_e^2 being the residual variance, and *r* being the number of blocks. The estimates for σ_g^2 and σ_e^2 were $[\text{EMS}(\text{G}) - \text{Var}(\text{Residual})]/r$ and $\text{Var}(\text{Residual})$, respectively. EMS(G) and Var (Residual) were obtained from the ANOVA table.

Genotyping

The common bean core set was genotyped with the BARCBean6K_3 Infinium BeadChips (Song et al., 2015), consisting of 5,398 SNPs distributed across the 11 pairs of common bean chromosomes with the Illumina BeadStation 500G (Gepts et al., 2019; Kuzay et al., 2020). The 5,389 SNPs across 382 accessions of the common bean core set are available and can be downloaded on the website at <https://datadryad.org/stash/dataset/10.25338/B8KP45>, with AA BB AB—format. The AA BB AB—was changed to the nucleotide format (A C T G) based on *P. vulgaris* G19833 reference sequences. After elimination of the missing data, a total of 4,654 SNPs were used for genetic diversity, population structure analysis, and GWAS in this study with a missing rate <20%, heterogeneous <10%, and minor allele frequency (MAF) > 5%. The distribution of the 4,654 SNPs on the 11 chromosomes of the common bean is shown in **Supplementary Figure 2**.

Genetic Diversity and Population Structure Analysis

This collection was previously analyzed with simple-sequence repeats (SSRs) (McClean et al., 2012) and SNPs (Gepts et al., 2019; Kuzay et al., 2020) for their genetic diversity and population structure. They found mainly three or seven subpopulations in the core set. In this study, we repeat the genetic diversity and population structure in the 315 accessions from the core set. A model-based clustering method in the STRUCTURE 2.3.4 program (Pritchard et al., 2000) was used to infer the population structure of the common bean accessions based on the 4,654 SNPs. To identify the number of populations (*K*) capturing the major structure in the data, the burn-in period was set at 50,000, with the Markov Chain Monte Carlo iterations, and the run

length was set at 10,000 in an admixture model; correlated allele frequencies were assumed to be independent for each run (Lv et al., 2012). Ten runs were performed for each simulated value of *K*, ranging from 1 to 10. For each simulated *K*, the statistical value delta *K* was calculated, using the formula described by Evanno et al. (2005). The optimal *K* was determined, using Structure Harvester (Earl and Vonholdt, 2012; <http://taylor0.biology.ucla.edu/structureHarvester/>). Each common bean genotype was then assigned to a cluster (*Q*) based on the probability determined by the software that the genotype belonged in the cluster. The cutoff probability for assignment to a cluster was 0.50 or above. Based on the optimum *K*, a bar plot with “Sort by *Q*” was obtained to show the population structure among the common bean genotypes (accessions).

The number of principal components (PC) was chosen according to the optimum subpopulation determined in STRUCTURE 2.3.4, and a PCA plot was drawn, using R package ggplot2 by the data from TASSEL 5 (Bradbury et al., 2007; <http://www.maizogenetics.net/tassel>). Genetic diversity also was assessed, and phylogenetic trees were drawn, using MEGA 7 (Kumar et al., 2016) based on the Maximum Likelihood (ML) tree method with the following parameters (Shi et al., 2016, 2017): the bootstrap method with the number of bootstrap replications 500; model/method: the General Time Reversible model; rates among sites: Gamma distributed with Invariant sites (*G* + *I*); the number of discrete gamma categories: five; gaps/missing data treatment: Use all sites; the ML heuristic method: Subtree-Pruning-Regrafting-Extensive (SPR level 5); the initial tree for ML: Make the initial tree automatically (Neighbor-Joining); and a branch swap filter: Moderate. During the drawing of the phylogeny trees, the population structure and the cluster information were imported to MEGA 7 for combined analysis of genetic diversity. For the sub-tree of each *Q* (cluster), the shapes of “Node/Subtree Marker” and the “Branch Line” were drawn with the same color as in the figure of the bar plot of the population clusters from the STRUCTURE 2.3.4 analysis.

Association Analysis

GWAS was performed, using the Genomic Association and Prediction Integrated Tool version 3 (GAPIT3) (Lipka et al., 2012; Wang and Zhang, 2020; <https://zzlab.net/GAPIT/index.html>; <https://github.com/jiabowang/GAPIT3>), where the mixed linear model (MLM), compressed MLM (CMLM) (Zhang et al., 2010), GLM, SUPER (Wang et al., 2014), multiple-locus MLM (MLMM), Fixed and Random Model Circulating Probability Unification (FarmCPU) (Liu et al., 2016), and Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) (Huang et al., 2019) were run in this study. Single marker regression (SMR), GLM (*Q*), and MLM (*Q*+*K*) were also conducted, using TASSEL 5 (Bradbury et al., 2007; <http://www.maizogenetics.net/tassel>). *Q*-matrix (*Q*) was obtained from the population structure analysis by STRUCTURE 2.3.4, and Kinship (*K*) was estimated by the tool Kinship with the Scald_IBS method built-in TASSEL 5. In addition, a *t*-test was performed for every single SNP, using visual basic codes in Microsoft Excel 2016. Multiple modes in several tools were used to identify SNP markers associated with resistance to SCN

HG Types to recognize more sTable NP markers and to tag the candidate gene(s) or QTL region(s) strongly associated with the SCN resistance. Highly significant associations were determined, using a strict Bonferroni correction of P -value at an $\alpha = 0.05$, in which the $P = 0.05/(\text{SNP number})$ as the significance threshold (López-Hernández and Cortés, 2019). In this study, for the panel of all 315 accessions, the significant LOD [$-\log_{10}(P\text{-value})$] [LOD was used instead of $-\log_{10}(P\text{-value})$ in the text] threshold value was 4.97, 4.84, and 4.52 for the panel of all 315 accessions, Q1, and Q2, respectively, based on the 4,654 SNPs, 3,455 SNPs, and 1,653 SNPs used for each panel after filtered with a missing rate $<20\%$, heterogeneity $<10\%$, and MAF $> 5\%$.

Besides the SCN phenotypic data of resistance to HG Type 0 (race 6) in the USDA common bean core collection from the experiment used to conduct GWAS for SCN resistance, the phenotypic data of resistance to HG Types 2.5.7 and 1.2.3.5.6.7 from the report by Wen et al. (2019) were also used to conduct GWAS, using the same BARCBean6K_3 Infinium BeadChips (Song et al., 2015). Although Wen et al. (2019) conducted the GWAS for the two HG Types, using 84,416 SNPs identified from GBS, more information and more SNP markers would be provided that are associated with resistance to HG Types 2.5.7 and 1.2.3.5.6.7 when using different SNP sets and different GWAS models. An LD heat map was drawn for regions containing a significant SNP marker, using Haploview (Barrett et al., 2005; <https://www.broadinstitute.org/haploview/haploview>). However, we do not conduct an LD-based haplotype association analysis in this research.

Candidate Gene Prediction

Candidate gene models were searched within 50 kb on either side of significant SNPs (Zhang H. Y. et al., 2016) and retrieved from the reference annotation of the common bean genome reference *Pvulgaris* v1.0_218 (<https://genome.jgi.doe.gov/portal/pages/dynamicOrganismDownload.jsf?organism=Phytozome>) because the SNP information was based on this reference sequence (Gepts et al., 2019).

Genomic Prediction of SCN Resistance

In this study, the ridge regression best linear unbiased prediction (RR-BLUP) was used to predict GEBV in GP and performed in the rrBLUP package (Endelman, 2011), with the R software Version 4 (<https://cran.r-project.org/bin/windows/base/rtest.html>). The RR-BLUP is an effective and accurate prediction method as demonstrated in a wide range of traits and crops (Jarquin et al., 2014; Zhang J. P. et al., 2016). In additions, GP was performed with the genomic best linear unbiased prediction (gBLUP) (Wang and Zhang, 2020; <https://zzlab.net/GAPIT/index.html>; <https://github.com/jiabowang/GAPIT3>) and also performed using Bayesian models: Bayes A, Bayes B, Bayes LASSO (BL), and Bayes ridge regression (BRR) (Legarra et al., 2011; Barili et al., 2018), random forest (RF) (Ogut et al., 2011), and support vector machines (SVM) (Maenhout et al., 2007). The “Bayesian Generalized Linear Regression (BGLR),” “RF,” and “kernlab” were used and run in the R package to perform the GP

models for Bayes A, Bayes B, BL, BRR, RF, and SVM (Bao et al., 2014; Ravelombola et al., 2019, 2020, 2021).

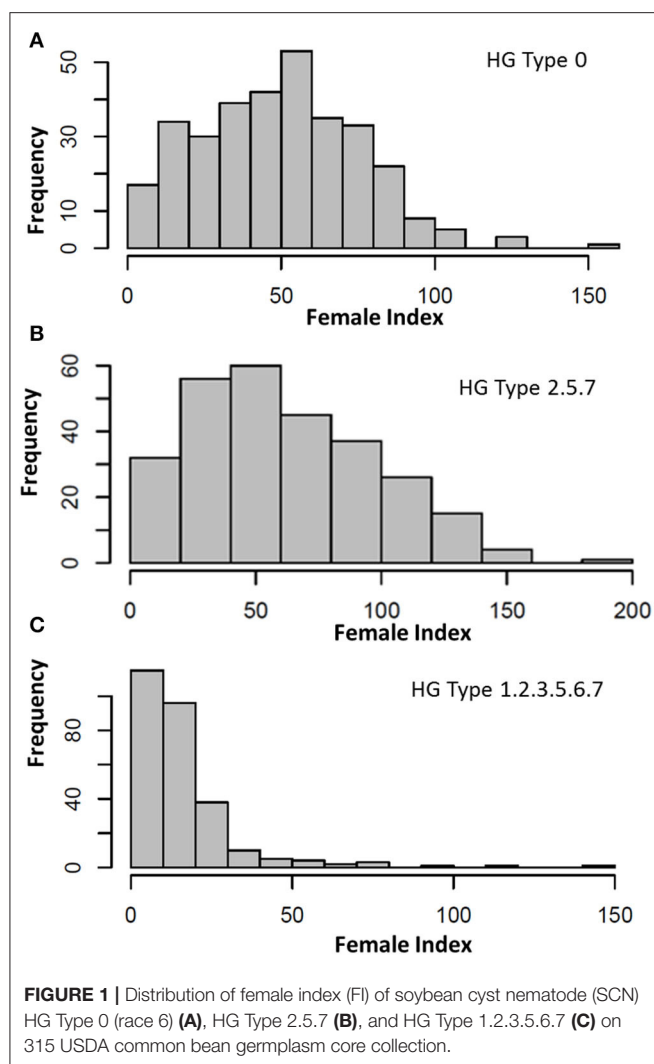
In this study, we conducted four groups of GP analyses (Bao et al., 2014; Tan et al., 2017; Ravelombola et al., 2019, 2020, 2021). (1) GP was performed with six different ratios of a training set: a testing set with 19:1, 9:1, 4:1, 2:1, and 1:1; or as 5, 10, 20, 30, 40, and 50% of a testing set in the panel of 315 common accessions. Each training population subset was randomly selected from the association panel, and the remainder was used as a testing set. (2) Nine different SNP number sets from 20 SNPs to all 4,654 SNPs were used in cross-predictions of resistance to three HG Types, using five GP models: rrBLUP, Bayes A, Bayes B, BL, and BRR. (3) Six different testing set sizes (percentages) from 5 to 50% were used in cross-prediction for resistance to three HG Types in three common bean populations (all tested accessions, Q1 population, and Q2 population), using a rrBLUP model. (4) Three SNP sets (all 4,654 SNPs, 20 SNP markers, and 20 random SNPs) were used in cross-prediction of resistance to three HG Types, using eight GP models (rrBLUP, gBLUP, Bayes A, Bayes B, BL, BRR, RF, and SVM). The PA was estimated using the average Pearson's correlation coefficient (r) between the GEBVs and observed phenotypic values for SCN resistance in the validation set (testing set) (Zhang J. P. et al., 2016; Shikha et al., 2017). The r -value indicates PA and the selection efficiency of GP; the higher the r -value, the more PA and the better the selection efficiency in GS. The training and testing sets were randomly created 100 times, and the r -value was estimated each time. The average r -value of 100 times was calculated for each trait (here for SCN HG Type 0, 2.5.7, or 1.2.3.5.6.7). The distribution charts were drawn by Microsoft Excel 2016 and R package ggplot2.

RESULTS

Soybean Cyst Nematode Resistance Evaluation

The reactions of SCN HG Type indicator lines and race differential lines to the SCN population are presented in **Supplementary Table 2**. In the HG Type test, the susceptible control Williams 82 soybean yielded 289 averaged SCN females per plant, indicating there was adequate SCN reproduction for this study. All of the seven HG Type indicators were resistant with FI < 10 , confirming that the SCN used in this study was the HG Type 0 (**Supplementary Table 2**). In the race test, the susceptible control Williams 82 soybean yielded 426 averaged females per plant, indicating there was adequate reproduction for this study. The lines PI 548982 (Pickett 71) and PI 548988 (Pickett) were moderately resistant to the SCN population with FI 19.3 and 25.6, respectively; and other indicator lines were resistant with FI < 10 , confirming the population was race 6 (**Supplementary Table 2**).

The FI of the HG Type 0 (race 6) on the common bean core accessions had a large range (145.5) from 4.8 on PI 313733 to 150.3 on PI 313671 (**Supplementary Tables 1, 3, Figure 1**), with an average of 49.9, SD of 25.45, SE of 1.43, and CV of 51.0%; and a near-normal distribution (**Figure 1A**), indicating a large variation of resistance reactions to the SCN HG Type 0. Fifteen



accessions were resistant to the HG Type 0 with $FI < 10$. The top seven accessions with the highest resistance to HG Type 0 were PI313733, PI201329, PI319684, PI313440, PI325614, PI417616, and PI313445, with FI ranging from 4.8 to 6.7, and the two most susceptible accessions were PI313671 with FI 150.3 and PI182004 with FI 124.5 (**Supplementary Tables 1, 3**). The H was 65.7%, indicating the HG Type 0 resistance was highly inheritable.

The FI of HG Type 2.5.7 ranged (199.1) from 0.4 on PI 313445 to 199.6 on PI 313671 (**Supplementary Tables 1, 3, Figure 1**), with an average of 62.9, SD of 36.4, SE of 2.19, and CV of 50.1%; a skewed near-normal distribution (**Figure 1**, middle) indicated a large variation of resistance reactions to this SCN HG Type. Twelve accessions were resistant to the HG Type 2.5.7 with $FI < 10$. The top seven accessions with the highest resistance were PI313445, PI417754, PI430210, PI201354, PI415913, PI417616, and PI325653, with FI ranged from 0.4 to 4.0; the two most susceptible accessions were PI313671, with FI 199.6 and PI 307820, with FI 158.6 (**Supplementary Tables 1, 3**).

The FI of HG Type 1.2.3.5.6.7 had a large range (146.1) from 0 for five accessions to 146.2 for PI 207148 (**Supplementary Tables 1, 3, Figure 1**), with an average of 15.9; SD of 17.0; a skewed distribution (**Figure 1**, bottom) indicated

there was a large variation of resistance reactions to this SCN HG Type. Fifty-nine out of the 276 accessions (21.4%) had $FI < 5.0$, and 115 out of 276 accessions (41.7%) had $FI < 10$, indicating there was a high percentage for the accessions resistant to the HG Type 1.2.3.5.6.7 (**Supplementary Table 1**). Many accessions were classified as resistant or highly resistant to HG Type 1.2.3.5.6.7, and only eight were susceptible ($FI > 65$). The two highest susceptible entries were PI207148 with FI 111.4 and PI313671 with FI 146.2.

Combining analysis of resistance to the three HG Types, only one accession, PI 313671, was susceptible with high $FI > 100$ for the three HG Types, indicating this accession can serve as a susceptible control. Four accessions, namely, PI201354, PI 313445, PI417616, and PI313733, had $FI < 10$ for resistance to the three HG Types, suggesting they have high and broad resistance to the three HG Types 0, 2.5.7, and 1.2.3.5.6.7 (**Supplementary Table 1**). There were 37 accessions with resistance to the three HG Types ($FI < 20$; **Table 1**); their genetic diversity will be analyzed in the following section of this report.

There were weak correlations ($r = 0.31$ to 0.33) of SCN resistance to HG Types, 0, 2.5.7, and 1.2.3.5.6.7 resistance among the 315 common bean accessions (**Supplementary Table 4**), suggesting that they had different genetic resistance to the three HG types.

From the 86 common bean accessions reported by Jain et al. (2019), 59 accessions were also screened for their resistance to HG Type 0 in this study. Six out of the 59 lines, PI201354, PI201329, PI430206, PI319684, PI343950, and PI269209, showed HG Type 0 resistance with $FI < 10$ in both studies, indicating the six lines had more durable or stable resistance. However, the correlation of the SCN HG Type resistance in the 59 lines between the two studies was very low, with $r = 0.057$, indicating that the SCN pathogens used in the two studies might have different pathogenicity. It is possible that the HG Type 0 population used in Jain et al. (2019) and the population we used belonged to different races because HG Type 0 can be race 3 or 6, and the race of the former was not reported.

Genetic Diversity and Population Structure Analysis

The population structure of the 315 USDA germplasm accessions was initially inferred, using STRUCTURE 2.3.4 (Pritchard et al., 2000). The peak delta K was observed at $K = 2$, indicating the presence of two main population clusters, Q1 and Q2, in the common bean germplasm panel (**Supplementary Figures 3A,B**). The classification of accessions into populations or clusters based on the model-based structure from STRUCTURE 2.3.4 is shown in **Supplementary Figure 3B** and **Supplementary Table 1**. The 315 accessions were assigned to one of the two populations or clusters, defined as Q1 and Q2 groups (populations). Q1 and Q2 consisted of 248 (78.7%) and 67 (21.3%) accessions, respectively (**Supplementary Table 1**). Seven accessions were classified as Q1(2) because their probability belonging to Q1 was >0.5 but <0.7 (**Supplementary Table 1**, bottom). A graphical plot of the PCA of the 315 common bean accessions showed two clusters (**Supplementary Figure 3C**) based on data from TASSEL 5 with two subpopulations.

TABLE 1 | Accession ID, origin (country), population clusters and groups, and their SCN Female Index (FI) of top 37 SCN-resistant common bean accessions in reaction to HG Types 0, 2.5.7 and 1.2.3.5.6.7.

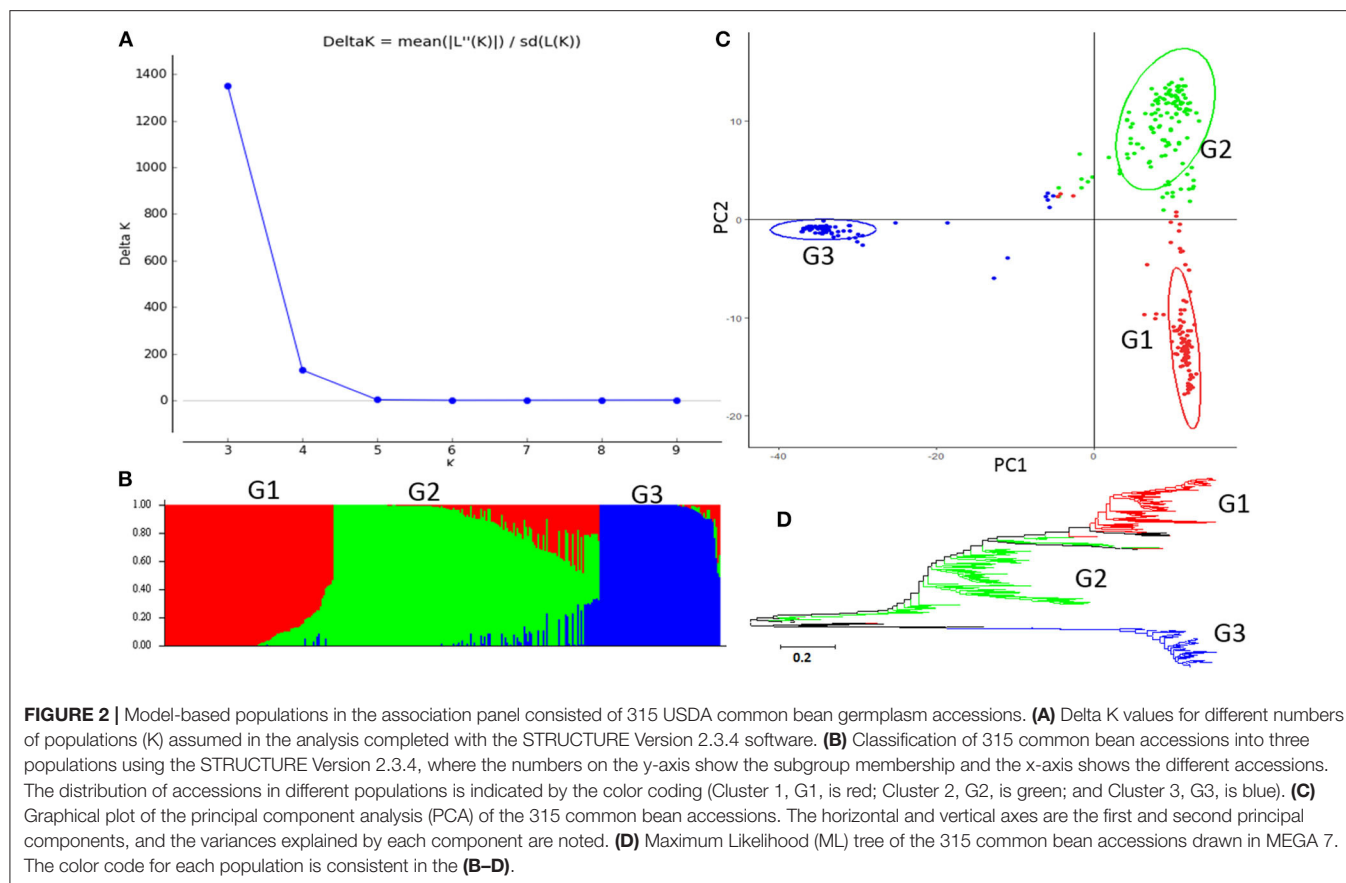
| Line_ID ^a | Line | Country | 2Q_cluster | 2_group | HG_Type 0_FI | HG_Type 2.5.7_FI | HG_Type 1.2.3.5.6.7_FI |
|--|----------|------------|------------|---------|--------------|------------------|------------------------|
| PI313615b_Colombia_Q1_0.987_0.013 ^a | PI313615 | Colombia | Q1 | I | 14.80 | | |
| PI313630b_Colombia_Q1_1_0 | PI313630 | Colombia | Q1 | I | 9.15 | | |
| PI309845_Costa Rica_Q1_1_0 | PI309845 | Costa Rica | Q1 | I | 11.66 | 28.19 | 4.49 |
| PI343950_Guatemala_Q1_1_0 | PI343950 | Guatemala | Q1 | I | 8.10 | | |
| PI449410_Mexico_Q1_1_0 | PI449410 | Mexico | Q1 | I | 14.54 | | |
| PI313328b_Mexico_Q1_1_0 | PI313328 | Mexico | Q1 | I | 7.02 | | |
| PI201329_Mexico_Q1_1_0 | PI201329 | Mexico | Q1 | I | 5.06 | 10.57 | 2.24 |
| PI201354_Mexico_Q1_1_0 | PI201354 | Mexico | Q1 | I | 7.19 | 3.08 | 0.37 |
| PI417667_Mexico_Q1_1_0 | PI417667 | Mexico | Q1 | I | 11.71 | 24.23 | 16.82 |
| PI313440_Mexico_Q1_1_0 | PI313440 | Mexico | Q1 | I | 5.92 | 8.81 | 17.2 |
| PI313445_Mexico_Q1_1_0 | PI313445 | Mexico | Q1 | I | 6.74 | 0.44 | 0.1 |
| PI313444_Mexico_Q1_1_0 | PI313444 | Mexico | Q1 | I | 7.08 | 16.74 | 10.28 |
| PI325630_Mexico_Q1_1_0 | PI325630 | Mexico | Q1 | I | 15.73 | 9.25 | 3.36 |
| PI417616_Mexico_Q1_1_0 | PI417616 | Mexico | Q1 | I | 6.46 | 3.96 | 7.29 |
| PI313473_Mexico_Q1_1_0 | PI313473 | Mexico | Q1 | I | 10.38 | | |
| PI203920_Mexico_Q1_1_0 | PI203920 | Mexico | Q1 | I | 19.41 | 25.55 | 12.71 |
| PI313501_Mexico_Q1_1_0 | PI313501 | Mexico | Q1 | I | 8.33 | 23.79 | 0.56 |
| PI325642_Mexico_Q1_1_0 | PI325642 | Mexico | Q1 | I | 11.22 | 10.13 | 3.74 |
| PI313512_Mexico_Q1_1_0 | PI313512 | Mexico | Q1 | I | 12.95 | 14.1 | 7.48 |
| PI201296_Mexico_Q1_1_0 | PI201296 | Mexico | Q1 | I | 14.11 | 12.78 | 1.87 |
| PI313490_Mexico_Q1_1_0 | PI313490 | Mexico | Q1 | I | 19.25 | 27.75 | 2.24 |
| PI325653_Mexico_Q1_1_0 | PI325653 | Mexico | Q1 | I | 16.21 | 3.96 | 1.5 |
| PI417739_Mexico_Q1_1_0 | PI417739 | Mexico | Q1 | I | 19.76 | 13.22 | 1.12 |
| PI430206_Mexico_Q1_1_0 | PI430206 | Mexico | Q1 | I | 9.40 | 12.33 | 0.1 |
| PI313820_Mexico_Q1_0.989_0.011 | PI313820 | Mexico | Q1 | I | 11.08 | | |
| PI313425_Mexico_Q1_1_0 | PI313425 | Mexico | Q1 | I | 15.09 | | |
| PI417657_Mexico_Q1_0.89_0.11 | PI417657 | Mexico | | I | 14.56 | 22.03 | 3.93 |
| PI430204_Mexico_Q1_0.692_0.308 | PI430204 | Mexico | | II(I) | 13.89 | 14.1 | 4.49 |
| PI346960_Mexico_Q1_0.661_0.339 | PI346960 | Mexico | Q1(2) | II(I) | 14.30 | 12.33 | 12.9 |
| PI345576_Costa Rica_Q1_0.672_0.328 | PI345576 | Costa Rica | Q1(2) | II(I) | 11.06 | 15.86 | 0.56 |
| PI241794_Ecuador_Q2_0.119_0.881 | PI241794 | Ecuador | Q2 | II | 14.55 | 20.7 | 15.89 |
| PI415936_Ecuador_Q2_0.027_0.973 | PI415936 | Ecuador | Q2 | II | 10.73 | 13.66 | 12.34 |
| PI209498_Costa Rica_Q2_0.019_0.981 | PI209498 | Costa Rica | Q2 | II | 11.47 | 28.19 | 17.01 |
| PI313733_Mexico_Q2_0_1 | PI313733 | Mexico | Q2 | II | 4.78 | 5.73 | 4.49 |
| PI325731_Mexico_Q2_0_1 | PI325731 | Mexico | Q2 | II | 17.58 | | |
| PI316030b_Peru_Q2_0_1 | PI316030 | Peru | Q2 | II | 13.51 | | |
| PI293355_Peru_Q2_0_1 | PI293355 | Peru | Q2 | II | 18.04 | 27.31 | 10.09 |

^aLine_ID consists of PI accession, original country, one of the two clusters Q1 or Q2, the Q1 probability, and Q2 probability. For example, PI313615b_Colombia_Q1_0.987_0.013, where the PI accession is PI313615b, which is grouped into Q1 cluster with probability of 0.987 and has 0.013 probability to Q2.

The genetic diversity among the 315 accessions was also assessed, using the ML method in MEGA 7 (Kumar et al., 2016), with phylogenetic trees are drawn based on the results. All accessions were assigned into one of the two clusters (populations), further indicating there were two distinct genetic populations in the common bean core set.

The second highest peak of delta K in STRUCTURE 2.3.4 was observed for K = 3, using Structure Harvester, indicating the 315 common bean germplasm accessions can be divided into three clusters (G1 to G3) (**Figure 2A**). **Figure 2B** shows

the bar plot drawn in STRUCTURE 2.3.4 to visualize the three-clustered populations. The classification of the germplasm accessions into populations based on the model-based structure developed in STRUCTURE 2.3.4 was shown in **Figure 2B**, **Supplementary Table 1**. Each common bean accession also was assigned to one of the three populations based on probabilities calculated in STRUCTURE 2.3.4 (**Supplementary Table 1**). A Q value = 0.5 was used to divide the three populations (clusters) and the admixture. In total, 301 out of 315 accessions (95.6%) were assigned to one of the three populations. G1



to G3 consisted of 97 (30.8%), 138 (43.8%), and 66 (21.0%) accessions, respectively. The remaining 14 accessions (4.4%) were categorized as having admixed ancestry between G1 and G3 (**Supplementary Table 1**). A PCA plot was shown in **Figure 2C** based on data from TASSEL 5.

The genetic diversity of the 315 common bean accessions was also analyzed, using the ML method in MEGA 7 by combining the three populations G1 to G3, identified by STRUCTURE. The results shown in **Figure 2D** indicate there may be three differentiated genetic populations and admixtures among the 315 accessions.

Combining (1) the two subpopulations (Q1 and Q2) and (2) the three clusters (G1 to G3) from STRUCTURE 2.3.4, a rectangular phylogenetic tree was drawn, using the ML method from MEGA 7 (**Supplementary Figure 4**). The common bean accession number, the original country of the accession, and the two populations (clusters) were merged into one taxon name for each branch in the combined tree drawn by MEGA 7 (**Supplementary Figure 4**). The resulting tree shows there were three main groups: (1) Q1G1, (2) Q1G2, and (3) Q2G3 in the 315 accessions (**Supplementary Figure 4**). Q1G1 included 96 accessions (30.5%), Q1G2 138 accessions (43.8%), Q1G (admixture) 8 (2.5%), Q2G3 66 (21.0%), Q2G31 (admixture) 1 (0.3%), and Q1(2) Gx (admixture) 7 (2.2%), indicating that the Q1 population was further divided into two groups and some admixture. The entire Q2 group (except

one) was not subdivided into the $K = 3$ analysis and became group G2 ($G2 \sim Q2 \sim Q2G2$), with only one exception (**Supplementary Table 1**), suggesting the Q2 population has a well-defined genetic background with stable boundaries.

Association Analysis

In this study, we performed GLM, MLM, SUPER, MLMM, FarmCPU, and BLINK analyses in GAPIT3 by setting $PCA = 3$, and SMR, GLM (Q), and MLM (Q+K) analyses in TASSEL 5, where $Q = 3$. We also conducted a t -test for each SNP. If an SNP had a LOD [$-\log(P\text{-value})$] greater than the significance threshold value LOD [$-\log(0.05/\text{SNP number})$] in one of the six MLM models (gapit.mlm, gapit.mlmm, gapit.super, gapit.farmCPU, gapit.blink, or tassell.mlm), the SNP was selected as a candidate-associated SNP marker and listed in **Supplementary Tables 5–7** for resistance to SCN HG Types 0, 2.5.7, and 1.2.3.5.6.7, respectively. After combining the output from GAPIT3 and Tassel 5 for the three association panels (all tested accessions (all.set), Q1 and Q2 populations), the SNP markers, which were significant for resistance to the three HG Types, are listed in **Table 2**.

Genome-Wide Association Study for Resistance to SCN HG Type 0

The distributions of the QQ plots between the observed *vs.* expected LOD [$-\log_{10}(p)$] showed a large divergence from

TABLE 2 | SNP markers associated with three SCN HG Types, 0, 2.5.7, and 1.2.3.5.6.7 in three sets of common bean genotypes, based on six models, BLINK, FarmCPU, MLM, MLMM, SUPER, and GLM in GAPIT 3 and three models, MLM, GLM, and SMR in Tassel 5, and *T*-test.

| SNP | Chr | Position | -log(P-value) using GAPIT 3 | | | | | | -Log(P-value) in Tassel | | | T-test | Rsquared in Tassel | | | R-allele | S-allele | MAF (%) | Set | Associated_HG_Type |
|-------------|-----|----------|-----------------------------|---------|------|------|-------|------|-------------------------|------|------|--------|--------------------|-------|-------|----------|----------|---------|---------|---------------------|
| | | | Blink | FarmCPU | MLM | MLMM | SUPER | GLM | SMR | GLM | MLM | | -LOG(P) | SMR | GLM | | | | | |
| ss715640464 | 4 | 33307678 | 10.31 | 4.31 | 4.90 | 1.08 | 0.79 | 6.27 | 0.03 | 5.74 | 4.16 | 0.38 | 0.05 | 7.47 | 6.33 | T | C | 22.6 | all.315 | HG Type 0 |
| ss715650114 | 6 | 10550456 | 5.68 | 5.11 | 2.60 | 2.70 | 2.36 | 2.91 | 3.04 | 2.32 | 1.97 | 3.38 | 4.39 | 3.09 | 2.94 | T | G | 26.6 | | |
| ss715647158 | 7 | 7343812 | 6.78 | 3.31 | 4.24 | 3.31 | 1.19 | 5.07 | 7.15 | 5.56 | 4.40 | 5.72 | 10.01 | 7.20 | 6.70 | C | A | 8.7 | | |
| ss715649511 | 7 | 7759866 | 5.76 | 3.36 | 2.43 | 2.75 | 5.26 | 4.27 | 1.69 | 3.53 | 1.74 | 2.44 | 2.47 | 4.64 | 2.64 | G | A | 46.2 | | |
| ss715639339 | 9 | 12175377 | 1.84 | 0.48 | 2.07 | 2.82 | 5.26 | 4.75 | 0.33 | 4.50 | 1.46 | 0.79 | 0.49 | 5.89 | 2.20 | T | C | 33.2 | | |
| ss715647549 | 11 | 44651807 | 7.42 | 2.67 | 3.37 | 3.89 | 5.40 | 4.64 | 1.13 | 3.74 | 2.33 | 1.34 | 1.83 | 5.41 | 4.00 | C | T | 37.3 | | |
| ss715639339 | 9 | 12175377 | 3.62 | 0.89 | 2.98 | 2.10 | 5.09 | 0.89 | 9.20 | 6.04 | 2.40 | 11.28 | 16.30 | 10.03 | 4.75 | T | C | 11.8 | Q1 | |
| ss715647549 | 11 | 44651807 | 7.26 | 0.83 | 4.25 | 5.46 | 5.36 | 0.83 | 7.16 | 3.71 | 2.72 | 9.61 | 14.47 | 6.96 | 5.98 | C | T | 19.5 | | |
| ss715641893 | 2 | 10113375 | 0.82 | 1.26 | 1.72 | 1.74 | 5.28 | 5.52 | 5.90 | 1.01 | 1.35 | 6.54 | 8.26 | 0.92 | 1.49 | T | C | 43.7 | all.315 | HG Type 2.5.7 |
| ss715639285 | 2 | 33312585 | 5.92 | 4.29 | 2.58 | 2.65 | 4.08 | 6.64 | 6.42 | 1.86 | 1.64 | 8.14 | 10.26 | 2.86 | 2.80 | T | G | 38.4 | | |
| ss715645573 | 3 | 50143102 | 3.39 | 5.67 | 2.86 | 2.94 | 3.97 | 5.67 | 5.64 | 2.21 | 2.52 | 6.07 | 9.10 | 3.39 | 4.34 | C | T | 41.1 | | |
| ss715645642 | 9 | 33052539 | 0.15 | 0.05 | 1.25 | 1.26 | 5.32 | 5.62 | 6.05 | 1.19 | 0.97 | 6.41 | 8.45 | 1.14 | 0.96 | G | T | 48.6 | | |
| ss715650604 | 1 | 41625385 | 1.11 | 0.02 | 1.59 | 1.62 | 8.91 | 5.44 | 5.21 | 2.79 | 1.08 | 7.58 | 11.16 | 5.69 | 2.46 | G | A | 39.8 | Q1 | |
| ss715651021 | 1 | 41732173 | 1.34 | 2.48 | 2.05 | 2.10 | 9.10 | 5.73 | 5.56 | 3.32 | 1.50 | 8.09 | 11.84 | 6.72 | 3.44 | T | C | 37.4 | | |
| ss715647960 | 1 | 41789504 | 1.12 | 0.04 | 1.49 | 1.51 | 8.45 | 5.26 | 4.96 | 2.29 | 0.88 | 7.29 | 10.60 | 4.68 | 2.01 | G | A | 38.1 | | |
| ss715639285 | 2 | 33312585 | 1.58 | 0.47 | 2.47 | 2.55 | 5.73 | 5.68 | 5.44 | 2.24 | 1.67 | 7.48 | 11.56 | 4.56 | 3.84 | T | G | 49.5 | | |
| ss715640488 | 7 | 35740746 | 6.07 | 6.17 | 3.62 | 3.82 | 10.35 | 8.86 | 9.67 | 6.09 | 2.63 | 18.50 | 19.60 | 11.88 | 6.10 | T | C | 22.8 | | |
| ss715640389 | 9 | 12154448 | 0.68 | 1.00 | 3.05 | 3.18 | 7.66 | 8.31 | 9.83 | 6.84 | 2.73 | 30.85 | 18.19 | 11.70 | 4.86 | A | C | 11.2 | | |
| ss715639339 | 9 | 12175377 | 4.88 | 2.20 | 2.98 | 3.10 | 8.42 | 8.40 | 9.03 | 6.18 | 2.05 | 26.01 | 18.45 | 12.06 | 4.74 | T | C | 11.9 | | |
| ss715641522 | 11 | 13037340 | 0.31 | 0.06 | 1.06 | 1.07 | 7.20 | 4.34 | 4.60 | 2.18 | 0.94 | 5.55 | 8.39 | 3.31 | 1.23 | T | C | 31.1 | | |
| ss715647636 | 3 | 3963582 | 9.29 | 5.65 | 2.55 | 4.38 | 2.73 | 2.56 | 8.18 | 2.61 | 2.54 | 2.78 | 11.57 | 2.97 | 3.12 | C | T | 11.2 | all.315 | HG Type 1.2.3.5.6.7 |
| ss715647109 | 6 | 27257765 | 10.60 | 8.32 | 3.43 | 2.92 | 3.79 | 3.43 | 8.55 | 3.29 | 3.14 | 2.76 | 13.44 | 4.84 | 5.14 | C | T | 9.4 | | |
| ss715640509 | 10 | 2792311 | 10.60 | 5.95 | 6.19 | 6.80 | 1.44 | 6.19 | 0.12 | 7.32 | 6.76 | 0.43 | 0.04 | 9.40 | 9.98 | T | C | 13.0 | | |
| ss715639563 | 11 | 46491205 | 6.09 | 4.15 | 1.92 | 2.37 | 3.33 | 1.92 | 7.27 | 2.62 | 2.58 | 4.02 | 13.38 | 4.56 | 4.76 | A | G | 27.0 | | |
| ss715639563 | 11 | 46491205 | 4.11 | 4.11 | 4.03 | 4.30 | 1.78 | 4.35 | 6.06 | 6.55 | 5.64 | 1.52 | 14.98 | 15.60 | 16.27 | A | G | 9.7 | Q1 | |
| ss715639339 | 9 | 12175377 | 1.84 | 0.48 | 2.07 | 2.82 | 5.26 | 4.75 | 0.33 | 4.50 | 1.46 | 0.79 | 0.49 | 5.89 | 2.20 | T | C | 33.2 | all | HG Type 0 |
| | | | 3.62 | 0.89 | 2.98 | 2.10 | 5.09 | 0.89 | 9.20 | 6.04 | 2.40 | 11.28 | 16.30 | 10.03 | 4.75 | T | C | 11.8 | Q1 | |
| | | | 0.33 | 0.54 | 2.14 | 2.18 | 1.36 | 0.31 | 0.33 | 4.15 | 1.65 | 0.56 | 0.57 | 6.26 | 2.83 | T | C | 33.2 | all | HG Type 2.5.7 |
| | | | 4.88 | 2.20 | 2.98 | 3.10 | 8.42 | 8.40 | 9.03 | 6.18 | 2.05 | 26.01 | 18.45 | 12.06 | 4.74 | T | C | 11.9 | Q1 | |

the expected distribution (**Supplementary Figure 5**), indicating there were SNPs associated with the resistance to SCN HG Type 0 in the three association panels. The Manhattan plot showed there were a dozen SNPs with LOD value > 4.97 in all.set and Q1 (**Supplementary Figure 5**), and associated with SCN resistance to HG Type 0. Based on MLM models, a total of 18 SNPs, located on Pv 03, 04, 05, 06, 07, 08, 09, and 11 had LOD > 4.79 in all.set or > 4.84 in Q1 (**Supplementary Table 5**), associated with the resistance to SCN HG Type 0 (**Supplementary Table 5**). Among the six models, BLINK had the highest LOD values, and several SNP markers were observed in all.set and Q1 but not in Q2 (**Figure 3**).

There were several SNPs with LOD > 4.97 in all.set and > 4.84 in the Q1 population in the SMR and GLM models but not in the MLM model (**Supplementary Table 5**), indicating that there were significant SNP markers, but they were not strongly associated with SCN resistance based on the Tassel tool. However, there were several SNPs with a LOD score > 4.0 or 3.0 , indicating there were small-effect QTLs for SCN resistance (**Supplementary Table 5**). Based on *t*-tests, all 18 SNPs had LOD values > 2.0 ($P < 0.01$) either in all.set, Q1, or Q2 (**Supplementary Table 5**).

After combining, six SNP markers, ss715640464, ss715650114, ss715647158, ss715649511, ss715639339, and ss715647549, located on chromosomes Pv04, 06, 07, 07, 09, and 11, were associated with resistance to SCN HG Type 0 in all.set (**Table 2**). The two SNPs, ss715647158 and ss715649511, were located at 7,343,812 bp and 7,759,866 bp, respectively, on Pv07 based on the Pvulgaris v1.0_218 whole-genome reference sequences (**Table 2**), suggesting that there was a QTL on Pv07 for HG Type 0 resistance. The ss715639339 SNP at 12,175,377 bp on Pv09 and ss715647549 at 44,651,807 bp on Pv11 were observed in both all.set and Q1 for HG Type 0 resistance (**Table 2**), suggesting the presence of a QTL on each of the two chromosomes.

Genome-Wide Association Study for Resistance to SCN HG Type 2.5.7

The distributions of the QQ plots between the observed vs. expected LOD $[-\log_{10}(p)]$ showed a large divergence from the expected distribution (**Supplementary Figure 6**), suggesting there were SNPs associated with resistance to SCN HG Type 2.5.7 in the three association panels. The Manhattan plot showed there were a dozen SNPs with a LOD value > 4.97 in all.set (**Supplementary Figures 6A,B**) and Q1 (**Supplementary Figures 6C,D**) for resistance to HG Type 2.5.7. A total of 15 SNPs, located on chromosomes Pv01, 02, 03, 07, 09, and 11 had LOD > 4.79 in all.set, or > 4.84 in Q1 (**Supplementary Table 6**). Among the six models, SUPER had the highest LOD values, and several SNP markers had LOD values greater than the 4.97 significance threshold in all.set, and > 4.84 in Q1, but not in Q2 (**Figure 4**, **Supplementary Table 6**).

The TASSEL 5 analysis showed that there were several significant SNPs with a LOD score > 4.97 in all.set and > 4.84 in the Q1 population in the SMR and GLM models but not in the MLM model (**Supplementary Table 6**). Nevertheless, these markers were not strongly associated with SCN resistance.

However, there were several SNPs with a LOD score > 3.0 or 2.5 , suggesting there were QTLs for HG Type 2.5.7 resistance with a small effect (**Supplementary Table 6**). Based on *t*-tests, 14 of the 15 SNPs had a LOD value > 2.0 ($P < 0.01$) either in all.set, Q1, or Q2, (**Supplementary Table 6**), indicating that the 14 SNPs were associated with resistance to HG Type 2.5.7 at the $P = 0.01$ significance level.

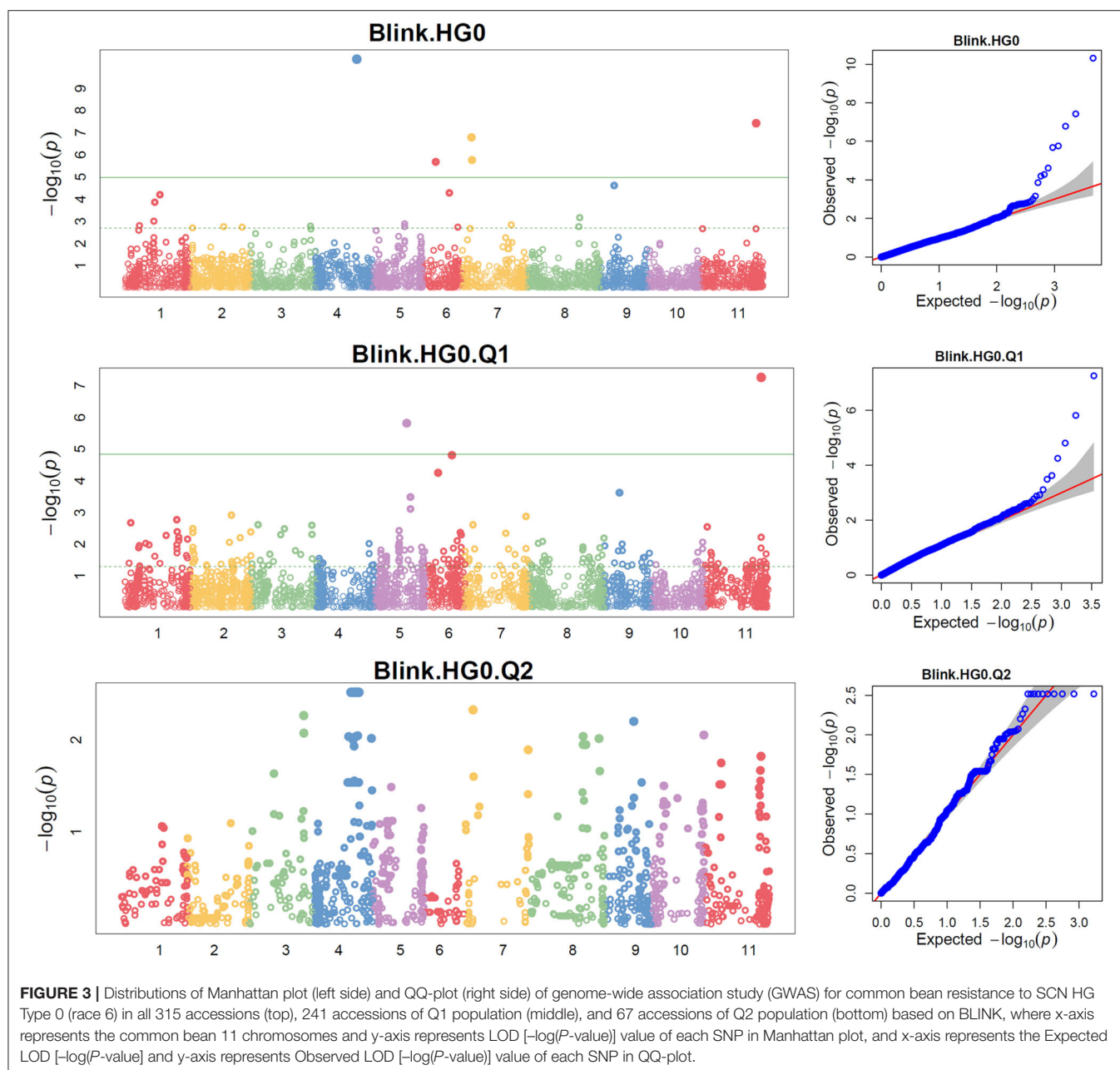
After combining, four SNPs were associated with resistance to the HG Type 2.5.7 in all.set, eight SNPs in Q1, and none in Q2 (**Table 2**). Among the eight SNPs in Q1, the three SNPs, ss715650604, ss715651021, and ss715647960, were located in the same region of chromosome Pv01, from 41,625,385 bp to 41,789,504 bp, indicating that there was a QTL on Pv01 for HG Type 2.5.7 resistance. The ss715639285 was identified in both all.set and Q1, suggesting that there was a QTL in the 33.3 Mbp region on Pv02 for HG Type 2.5.7 resistance. The two SNPs, ss715640389 and ss715639339, were located in the same region, from 12,154,448 bp to 12,175,377 bp on Pv09, and the two SNPs had very high LOD values (> 26) in the *t*-test (**Table 2**). In addition, a SNP, ss715640488 at 35,740,746 bp on Pv07 and another SNP, ss715641522, at 13,037,340 bp on Pv11 were associated with HG Type 2.5.7 resistance.

Genome-Wide Association Study for Resistance to SCN HG Type 1.2.3.5.6.7

The distributions of the QQ plots between the observed vs. expected LOD $[-\log_{10}(p)]$ showed a large divergence from the expected distribution (**Supplementary Figure 6**), indicating there were SNPs associated with resistance to SCN HG Type 1.2.3.5.6.7 in the three association panels. The Manhattan plot showed there were several SNPs with LOD values > 4.97 in all.set (**Supplementary Figures 7A,B**), suggesting there were SNPs associated with SCN resistance to HG Type 1.2.3.5.6.7. Six SNPs, ss715647636, ss715647109, ss715647614, ss715649401, ss715640509, and ss715639563, located on chromosomes Pv 03, 06, 09, 09, 10, and 11, respectively had LOD > 4.79 in all.set (**Supplementary Table 7**). Among the six models, BLINK had the highest LOD values, and several SNP markers were observed with a significant LOD value > 4.97 in all.set but not in Q1 or Q2 (**Figure 5**), indicating there were significant SNPs associated with SCN resistance to HG Type 1.2.3.5.6.7 based on the association panel of all.set of 276 accessions. Two additional SNPs, ss715646397 and ss715648134, located on Pv03 and 04, also had LOD values greater than four and were selected as markers for HG Type 1.2.3.5.6.7 resistance in the Q1 population (**Supplementary Table 7**).

There were only three SNPs that had a LOD score > 4.97 in all.set and one SNP with LOD > 4.84 in the Q1 population, either in SMR, GLM, or MLM models (**Supplementary Table 7**). However, seven out of the eight listed SNPs had LOD > 3.0 or 2.5 in all.set or Q1, suggesting there were QTLs for SCN resistance with small effects (**Supplementary Table 7**). The *t*-tests indicated that the eight SNPs had a LOD value > 2.0 ($P < 0.01$) either in all.set, Q1, or Q2 (**Supplementary Table 7**).

After combining, four SNPs were associated with resistance to SCN HG Type 1.2.3.5.6.7 in all.set, one SNP in Q1, and none in



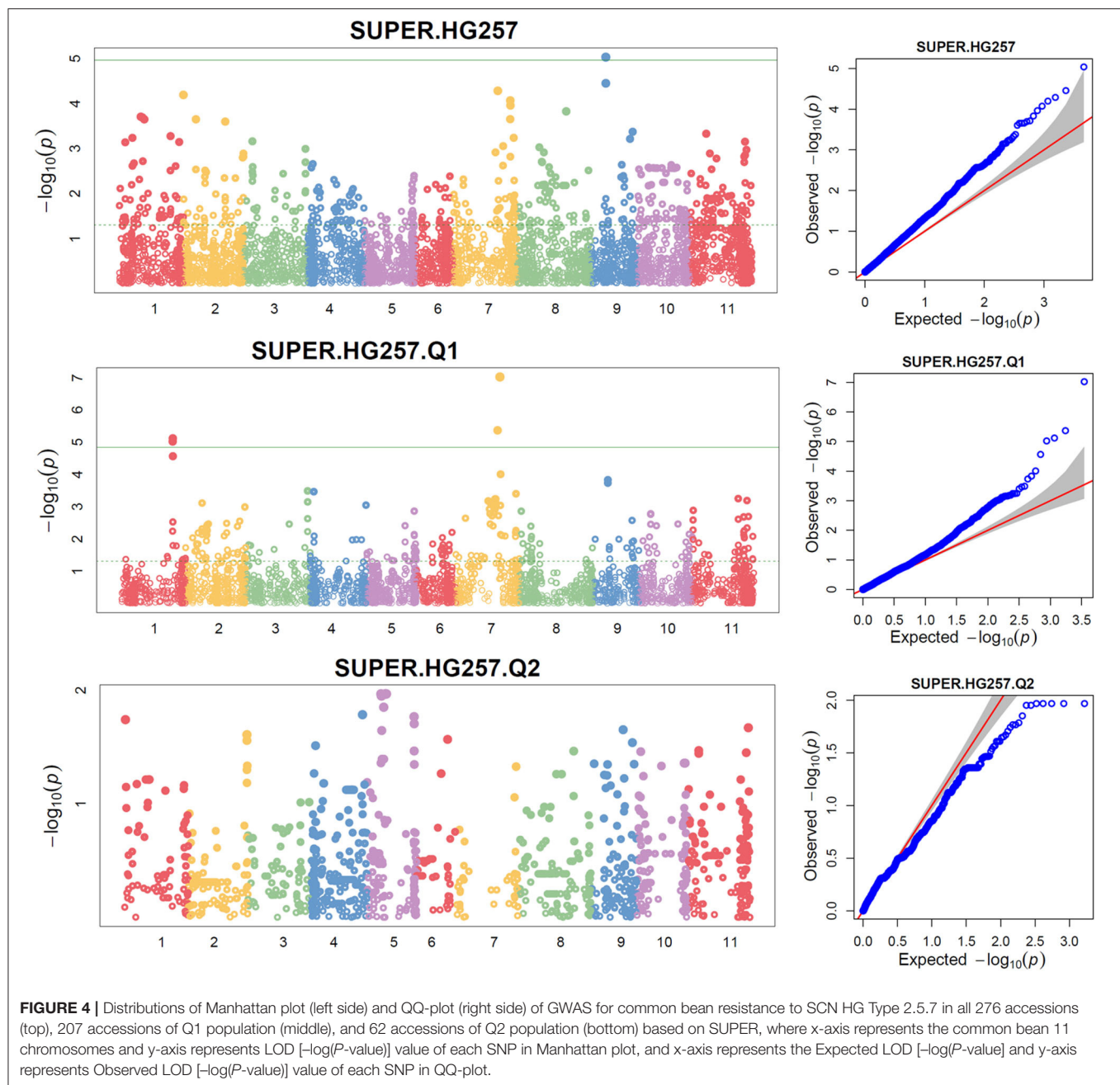
Q2 (Table 2). The four SNP markers in all.set were ss715647636, ss715647109, ss715640509, and ss715639563, located at 3,963,582 bp, 27,257,765 bp, 2,792,311 bp, and 46,491,205 bp on Pv 03, 06, 10, and 11, respectively (Table 2). SNP marker ss715639563 was also identified in Q1 population, increasing the confidence in this SNP as a marker for HG Type 1.2.3.5.6.7 resistance.

Combining GWAS for Resistance to the Three SCN HG Types

In this study, a total of 40 SNPs were identified as potential SNP markers associated with SCN resistance (Supplementary Tables 5–7) based on the LOD values from the MLM models in GAPIT3 and Tassel 5, after Bonferroni

correction. Combining results from the six models (GLM, MLM, SUPER, MLMM, FarmCPU, and BLINK) in GAPIT3, three models (SMR, GLM, and MLM) in TASSEL 5, and *t*-tests among the three association panels (all.set, Q1, and Q2), 6, 11, and 4 SNPs were significantly associated with resistance to HG Type 0, 2.5.7, and 1.2.3.5.6.7, respectively (Table 2). Among them, one SNP, ss715639339, at 12,175,377 bp on Pv09 was associated with the resistance to both SCN HG Types 0 and 2.5.7 (Table 2).

We did not conduct LD analysis for all SNPs in this study. However, the LD heatmaps were drawn, using Haploview for seven genome regions with the eight SNP markers significantly associated with resistance to either SCN HG Type 0, 2.5.7, or 1.2.3.5.6.7 (Supplementary Figure 8), where two LLR genes were

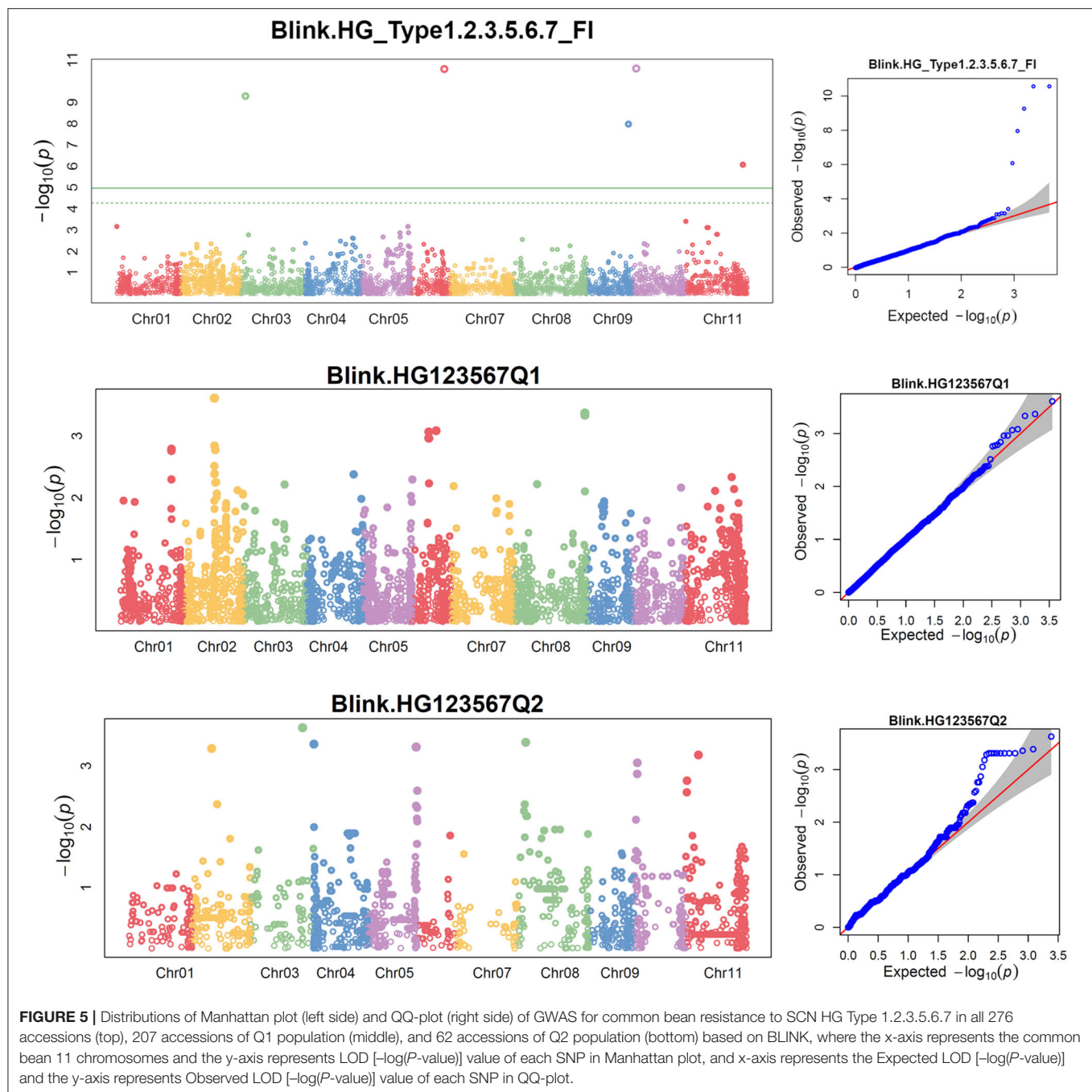


also included: Phvul.006G104700 and Phvul.010G018300. The Phvul.006G104700 gene is located on Pv04 in the same LD block as an SNP marker SS715640464 at a distance of only 8.98 Kbp (**Supplementary Figure 8A**) for HG Type 0 resistance. The gene Phvul.010G018300 is located on Pv10 at a distance of 39.9 Kbp from ss715640509 associated with HG Type 1.2.3.5.6.7 resistance (**Supplementary Figure 8E**, bottom left).

Candidate Genes for SCN Resistance

A total of 20 significant GWAS-derived SNPs were selected as markers associated with the resistance to the three SCN HG Types, 0, 2.5.7, and 1.2.3.5.6.7 (**Table 2**). Candidate gene

models were searched within 10, 30, and 50 kb, flanking each of these SNPs. A total of 125, 83, 33, 19, and 8 genes were found at a distance of 50, 30, 10, 5, and 1 Kb, respectively, from the 20 SNPs (**Supplementary Table 8**) based on the annotations of the common bean genome reference Pvulgaris v1.0_218. Among the 125 genes, five gene models, Phvul.001G158800, Phvul.002G072100, Phvul.006G160700, Phvul.007G080900, and Phvul.009G223200, contained an SNP marker, ss715647960, ss715641893, ss715647109, ss715649511, and ss715645642, respectively, on chromosomes Pv01, Pv02, Pv06, Pv07, and Pv09 (**Table 3**). Whether these five gene models are related to SCN resistance needs further study.



The Leucine-Rich Repeat (LRR) gene model Phvul.004G099300 (disease resistance family protein/LRR family protein), located at 33,316,658–33,320,257 bp on Pv04, based on the common bean genome reference *Pvulgaris* v1.0_218, is located near the SNP marker ss715640464 (distance of 8.98 Kbp), associated with SCN HG Type 0 resistance. Another LRR gene, Phvul.010G018300 (LRR protein kinase family protein) at 2,832,211–2,839,756 bp on Pv10 is close to the SNP marker ss715640509 (distance of 39.9 Kbp). In addition, one NAC-domain gene, Phvul.006G023100 (NAC-domain

containing protein 42), is located at 10,522,343–10,526,782 bp on Pv06 was close (~ 24 Kbp) to the SNP marker ss715650114 (Table 3). Whether the two LRR genes and the NAC-domain gene are related to SCN resistance needs further evaluation.

Genomic Prediction of SCN Resistance

Genomic Prediction With Different Ratios of a Training Set to a Testing Set

In this study, GP was performed using six different ratios of training/testing sets, 19:1, 9:1, 4:1, 7:3, and 1:1, expressed as 5,

TABLE 3 | Candidate genes for SCN resistance in common bean.

| Gene | Chr | Start | End | Arabi-defline | SNP | Chr | Position | Distance from the gene start | Distance from the gene end | Distance from the SNP to the gene | Associated SCN HG Type |
|--------------------|-----|----------|----------|--|-------------|-----|----------|------------------------------|----------------------------|-----------------------------------|------------------------|
| Phvul.001G1588001 | | 41789351 | 41790573 | C2H2 and C2HC zinc fingers superfamily protein | ss715647960 | 1 | 41789504 | 153 | -1,069 | on the gene | HG257 |
| Phvul.002G0721002 | | 10112472 | 10116406 | aldehyde dehydrogenase 2B7 | ss715641893 | 2 | 10113375 | 903 | -3,031 | on the gene | HG257 |
| Phvul.004G1047004 | | 33316658 | 33320257 | disease resistance family protein/LRR family protein | ss715640464 | 4 | 33307678 | -8,980 | -12,579 | within 10 kb distance | HG0 |
| Phvul.006G0231006 | | 10522343 | 10526782 | NAC domain containing protein 42 | ss715650114 | 6 | 10550456 | 28,113 | 23,674 | within 30 kb distance | HG0 |
| Phvul.006G1607006 | | 27256833 | 27258542 | sugar transporter 1 | ss715647109 | 6 | 27257765 | 932 | -777 | on the gene | HG123567 |
| Phvul.007G0809007 | | 7759099 | 7762514 | Protein of unknown function (DUF677) | ss715649511 | 7 | 7759866 | 767 | -2,648 | on the gene | HG0 |
| Phvul.009G2232009 | | 33050989 | 33058854 | ARID/BRIGHT DNA-binding domain-containing protein | ss715645642 | 9 | 33052539 | 1,550 | -6,315 | on the gene | HG257 |
| Phvul.010G01830010 | | 2832211 | 2839756 | Leucine-rich repeat protein kinase family protein | ss715640509 | 10 | 2792311 | -39,900 | -47,445 | within 50 kb distance | HG123567 |

10, 20, 30, 40, and 50% of a testing set in all.set, containing the 315 common bean accessions for HG Type 0 resistance or 276 accessions for HG Types 2.5.7 and 1.2.3.5.6.7 resistance. The actual sizes of the [training set/testing set] were 299/16, 283/32, 252/63, 220/95, 189/126, and 158/157 for HG Type 0, and 262/14, 248/28, 221/55, 193/83, 166/110, and 138/138 for HG Types 2.5.7 and 1.2.3.5.6.7. The GEBVs and r -values between GEBVs and observed values in the testing set were estimated by six GP models (rrBLUP, gBLUP, Bayes A, Bayes B, BL, and BRR) in cross-prediction for resistance to the three HG Types, 0, 2.5.7, and 1.2.3.5.6.7, using (1) all 4,654 SNPs and (2) 20 associated SNP markers (20 GWAS-derived SNP markers). There were six ratios between training and testing sets, six models, two SNP sets, and three SCN HG types to make a total of 216 combinations. Each combination was run 100 times to calculate GP statistical parameters and r -values. The average r -value ($r_{\bar{Y}_{100}}$) and its SE from the 100 runs for each combination are listed in Supplementary Table 9 and the 216 averaged r -values ($r_{\bar{Y}_{100}}$) displayed in charts drawn in MS Excel 2016, grouped by the six sets (5, 10, 20, 30, 40, and 50%) of testing set percentages (Supplementary Figure 9). The r -distribution charts were created by an R-package for the 216 combinations grouped by percentages of a testing set; the r -distributions of the 36 combinations estimated by rrBLUP model are listed in Figure 6. The 108 averaged r -values ($r_{\bar{Y}_{100}}$) (half of all 216 combinations) for the all.set are listed in Table 4.

The six sets of 5, 10, 20, 30, 40, and 50% of testing set percentages had similar, although not identical averaged r -values across five models except gBLUP with slightly lower r -values (Table 4, Figure 4, Supplementary Figure 9, Supplementary Table 9). The r -value, averaged over six models, was 0.39 for HG Type 0, 0.33 for HG Type 2.5.7, and 0.27 for HG Type 1.2.3.5.6.7. They were 0.40 for HG Type 0, 0.35 for HG Type 2.5.7, and 0.31 for HG Type 1.2.3.5.6.7 when averaged from five models, except gBLUP, when using all 4,654 SNPs (Table 4, Supplementary Table 9). This observation suggested that it may be feasible to do GS for SCN resistance in common bean with one of the six sets. The r -value increased to 0.46, 0.38, and 0.41, averaged over the six models, and 0.51, 0.41, and 0.46, averaged over the five models (except gBLUP) when using only the 20 SNP markers (Supplementary Table 9, Supplementary Figure 9), suggesting that GWAS-derived SNP markers can be used in GS. From Figure 6, the 5% test set had the largest variance, and the 50% test set had the smallest. The PA decreased when the size of the testing set increased. Likewise, the SE values decreased when the test sets increased from 5 to 50% (Supplementary Table 9), indicating that the larger the testing set, the less variable the r -values. However, a small decrease of the r -value was observed as well in most cases when the training/test ratio was 40% or higher.

Genomic Prediction With Different SNP Numbers

Genomic prediction was performed with nine different SNP number sets (20, 50, 100, 200, 400, 800, 1,600, and all 4,654 SNPs, plus the 20 GWAS-derived SNP markers) in cross-predictions for resistance to three HG Types, using five GP models: rrBLUP, Bayes A, Bayes B, BL, and BRR. There were 135

combinations for GP analysis, consisting of nine SNP sets, five GP models, and three SCN HG Types. Each combination was run 100 times to calculate GP statistical parameters and r-values. The average r-value ($r_{\bar{Y}_{100}}$) and its SE from the 100 runs for each combination are presented in the **Supplementary Table 10**, **Supplementary Figure 10**. The 27 averaged r-values ($r_{\bar{Y}_{100}}$) estimated by rrBLUP are presented in **Table 5**. The 54 r-distribution charts created by ggplots in R-package for r-values, estimated by Bayes A and rrBLUP models, are shown in **Figure 7**.

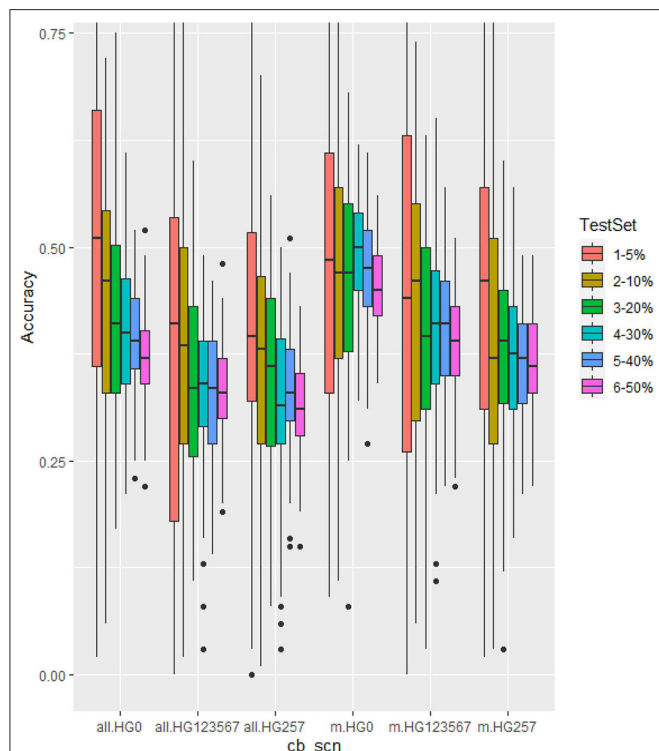


FIGURE 6 | Genomic prediction of six different testing set percentages from 5 to 50% in cross-prediction for resistance to three SCN HG Types, 0, 1.2.3.5.6.7, and 2.5.7 using all 4,654 SNPs (left three groups as all.HG0, all.HG123567, and all.HG257), and 20 associated SNP markers (m.HG0, m.HG123567, and m.HG257) estimated by rrBLUP model.

The nine SNP sets had an averaged r-value 0.38 for HG Type 0, 0.31 for HG Type 2.5.7, and 0.33 for HG Type 1.2.3.5.6.7 (**Table 5**, **Figure 7**, **Supplementary Figure 10**). The r-values were somewhat decreased a little when 100 or less SNPs were used for HG Type 0 resistance, 200 or less SNPs were used for HG Type 2.5.7 resistance, but did not decrease for HG Type 1.2.3.5.6.7 resistance, indicating that sets of more than 200 SNPs can be used for GS. The set of the 20 SNP markers had the highest averaged r-values in all five models for the three HG Type resistances (**Figure 7**), indicating that the 20 associated SNP markers can be used to do GS for SCN resistance selection as well.

Genomic Selection in Three Association Panels

Genomic prediction was performed in the three association panels (all.set, Q1, and Q2) with six different testing set sizes from 5 to 50% in cross-prediction for resistance to the three HG Types, using the rrBLUP model (54 combinations). Each combination was run 100 times to estimate GEBVs and r-values. The average r-value ($r_{\bar{Y}_{100}}$) and its SE from the 100 runs for each combination are listed in **Supplementary Table 11**, and the r-charts are shown in **Supplementary Figure 11**.

For the HG Type 0 resistance, all r-values are similar among the three sets (all.set, Q1, and Q2) across six testing sets with averaged 0.41, 0.41, and 0.38, respectively (**Supplementary Table 11**, top). For HG Type 2.5.7 and 1.2.3.5.6.7 resistance, all.set and Q1 had similar r-values, but Q2 had much lower r-values (**Supplementary Figure 11**). The 5% of the “Testing set” had the largest variability, and the 50% had the lowest SE value, and PA decreased when the “Testing set” percentage increased (**Supplementary Table 11**).

Genomic Prediction Comparisons Among All SNPs, SNP Markers, and the Random SNP Set

Genomic prediction was performed for three SNP sets (all 4,654 SNPs, 20 GWAS-derived SNP markers, and 20 random SNPs) in cross-prediction for resistance to three HG Types, using eight GP models (rrBLUP, gBLUP, Bayes A, Bayes B, BL, BRR, RF, and SVM) (72 combinations). Each combination was run 100 times to estimate GEBVs and r-values. The average r-value ($r_{\bar{Y}_{100}}$) and SE from the 100 runs for each combination are presented in **Supplementary Table 12**, and the r-charts are also showed in **Supplementary Figure 12**.

TABLE 4 | Prediction accuracy (PA) for SCN resistance to three HG Types with six different testing sets (percentages) using all 4,654 SNPs with six genomic prediction models.

| GP model | r-value in HG Type 0 | | | | | | | r-value in HG Type 2.5.7 | | | | | | | r-value in HG Type 1.2.3.5.6.7 | | | | | | |
|----------|----------------------|------|------|------|------|------|---------|--------------------------|------|------|------|------|------|---------|--------------------------------|------|------|------|------|------|---------|
| | 5% | 10% | 20% | 30% | 40% | 50% | Average | 5% | 10% | 20% | 30% | 40% | 50% | Average | 5% | 10% | 20% | 30% | 40% | 50% | Average |
| rrBLUP | 0.44 | 0.41 | 0.41 | 0.41 | 0.40 | 0.37 | 0.41 | 0.33 | 0.36 | 0.35 | 0.32 | 0.33 | 0.32 | 0.34 | 0.30 | 0.36 | 0.34 | 0.33 | 0.33 | 0.33 | 0.33 |
| gBLUP | 0.38 | 0.31 | 0.30 | 0.29 | 0.28 | 0.27 | 0.31 | 0.25 | 0.31 | 0.27 | 0.26 | 0.24 | 0.23 | 0.26 | 0.11 | 0.11 | 0.12 | 0.10 | 0.08 | 0.08 | 0.10 |
| Bayes A | 0.41 | 0.39 | 0.39 | 0.42 | 0.39 | 0.39 | 0.40 | 0.33 | 0.40 | 0.37 | 0.36 | 0.35 | 0.34 | 0.36 | 0.31 | 0.31 | 0.30 | 0.30 | 0.29 | 0.29 | 0.30 |
| Bayes B | 0.40 | 0.40 | 0.38 | 0.40 | 0.37 | 0.36 | 0.39 | 0.34 | 0.35 | 0.33 | 0.35 | 0.33 | 0.31 | 0.33 | 0.32 | 0.30 | 0.30 | 0.31 | 0.29 | 0.28 | 0.30 |
| BL | 0.43 | 0.43 | 0.43 | 0.40 | 0.38 | 0.38 | 0.41 | 0.37 | 0.33 | 0.35 | 0.37 | 0.35 | 0.34 | 0.35 | 0.27 | 0.34 | 0.28 | 0.29 | 0.27 | 0.28 | 0.29 |
| BRR | 0.44 | 0.41 | 0.40 | 0.41 | 0.39 | 0.38 | 0.41 | 0.38 | 0.36 | 0.38 | 0.35 | 0.34 | 0.32 | 0.36 | 0.31 | 0.35 | 0.34 | 0.34 | 0.31 | 0.31 | 0.33 |
| Average | 0.42 | 0.39 | 0.39 | 0.39 | 0.37 | 0.36 | 0.39 | 0.33 | 0.35 | 0.34 | 0.34 | 0.32 | 0.31 | 0.33 | 0.27 | 0.30 | 0.28 | 0.28 | 0.26 | 0.26 | 0.27 |

TABLE 5 | Genomic prediction of nine different SNP number sets from 20 SNPs to all 4,654 SNPs in cross-prediction for resistance to SCN HG Types 0, HG Type 2.5.7, and HG Type 1.2.3.5.6.7 using rrBLUP.

| HG Type | 4654SNP | 1600SNP | 800SNP | 400SNP | 200SNP | 100SNP | 50SNP | 20SNP | 20SNP.marker | Average |
|----------------|---------|---------|--------|--------|--------|--------|-------|-------|--------------|---------|
| HG Type 0 | 0.41 | 0.38 | 0.44 | 0.42 | 0.45 | 0.33 | 0.22 | 0.27 | 0.52 | 0.38 |
| HG Type 257 | 0.35 | 0.34 | 0.32 | 0.34 | 0.30 | 0.28 | 0.22 | 0.21 | 0.40 | 0.31 |
| HG Type 123567 | 0.34 | 0.34 | 0.30 | 0.34 | 0.33 | 0.32 | 0.31 | 0.34 | 0.39 | 0.33 |

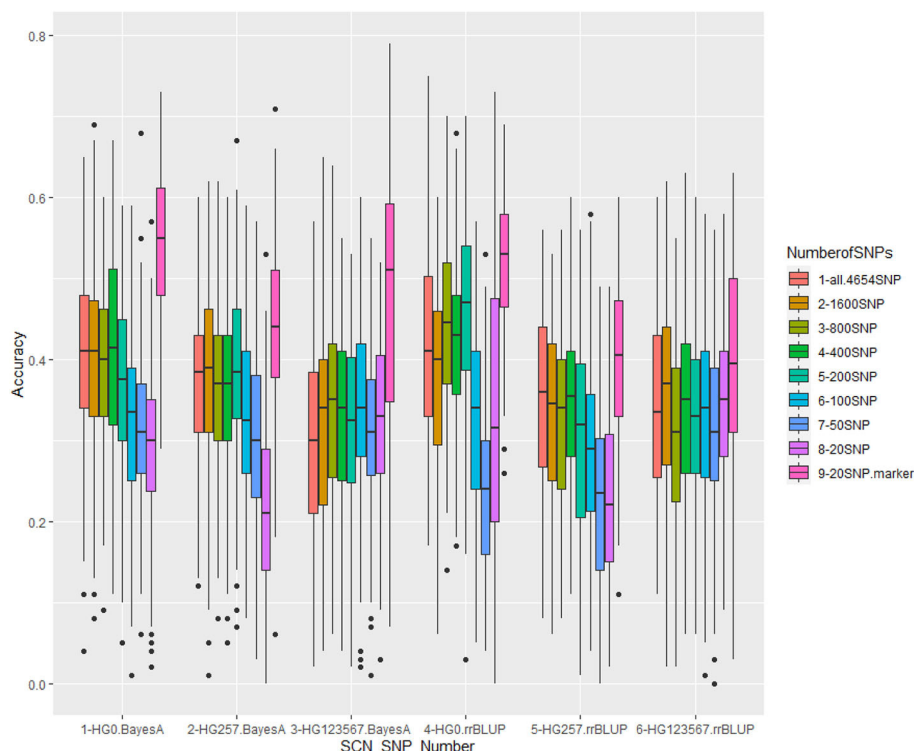


FIGURE 7 | Genomic prediction of nine different SNP numbers from 20 SNPs to all 4,654 SNPs in cross-prediction for resistance to three SCN HG Types, 0, 2.5.7, and 1.2.3.5.6.7 using Bayes A model (left three groups) and rrBLUP model (right three groups).

The set of 20 GWAS-derived SNP markers had the highest *r*-values across the eight models for resistance to either HG Type 0, 2.5.6, or 1.2.3.5.6.7, suggesting that the GWAS-derived SNP markers will be more effective for GS than the random 20-SNP sets (**Supplementary Table 12, Supplementary Figure 12**). The set of “random 20 SNPs” had the lowest *r*-values, suggesting that using more SNPs would increase the selection effectiveness in GS.

Genomic Prediction Using Different Models

Eight GP models (rrBLUP, gBLUP, Bayes A, Bayes B, BL, BRR, RF, and SVM) were used to conduct GP for resistance to the three HG Types. The five GP models (rrBLUP, Bayes A, Bayes B, BL, and BRR) had similar *r*-values, but the gBLUP model had the lowest *r*-values for resistance to either HG Type 0, 2.5.7, or 1.2.3.5.6.7 (**Supplementary Figure 13C**).

Based on the results from six different testing sets (percentages) in 315 common bean accessions (**Table 4, Supplementary Table 9**), the six GP models (rrBLUP, gBLUP,

Bayes A, Bayes B, BL, and BRR) had similar averaged PA (0.41, 0.31, 0.40, 0.39, 0.41, and 0.41) for resistance to HG Type 0; lower but similar PA (0.34, 0.26, 0.36, 0.33, 0.35, and 0.36) for HG Type 2.5.7 resistance; and the lowest PA (0.33, 0.10, 0.30, 0.30, 0.29, and 0.33) for HG Type 1.2.3.5.6.7 resistance. When the set of 20 significant SNP markers was used, the averaged PA of the six models increased for resistance to all of the three HG Types (**Supplementary Table 9, bottom half**).

Based on the results from the nine different SNP number sets from 20 SNPs to all 4,654 SNPs in cross-prediction for resistance to the three HG Types (**Supplementary Table 10, Supplementary Figure 11**), the five GP models (rrBLUP, Bayes A, Bayes B, BL, and BRR) had averaged PA, 0.38, 0.38, 0.36, 0.38, and 0.38, respectively, for resistance to HG Type 0; 0.31, 0.35, 0.31, 0.34, and 0.35 for HG Type 2.5.7 resistance; and 0.33, 0.34, 0.30, 0.34, and 0.34 for HG Type 1.2.3.5.6.7 resistance.

Based on the three SNP sets (all 4,654 SNPs, 20 significant SNP markers, and 20 random SNPs) used in cross-prediction,

the eight GP models, rrBLUP, gBLUP, Bayes A, Bayes B, BL, BRR, RE, and SVM, had averaged PA values of 0.38, 0.20, 0.41, 0.35, 0.42, 0.40, 0.35, and 0.39, respectively, for resistance to HG Type 0; 0.31, 0.20, 0.34, 0.28, 0.34, 0.34, 0.29, and 0.32 for HG Type 2.5.7 resistance; and 0.36, 0.11, 0.37, 0.32, 0.38, 0.39, 0.33, and 0.30 for HG Type 1.2.3.5.6.7 resistance (**Supplementary Table 12**, **Supplementary Figure 12**).

Overall, sets of 400 SNPs or more for GP had similar GS efficiency (r-value) for resistance to either HG Type 0, 2.5.7, or 1.2.3.5.6.7. The set of 20 significant SNP markers for GP had the highest r-value for GP (**Supplementary Figure 13A**). The six sets of different sizes from 5 to 50% had similar r-values (**Supplementary Figure 13B**). Except for the gBLUP model, which had a lower r-value for GP, all other seven models had similar PA (**Supplementary Figure 13C**). The averaged r-values were 0.40 for HG Type 0 resistance, 0.34 for HG Type 2.5.7, and 0.32 for HG Type 1.2.3.5.6.7 (**Supplementary Figure 13D**), indicating that we can use one of the seven GP models to conduct GS. Each model provided similar selection efficiency for SCN resistance.

Genomic Heritability (GH)

In this study, the GH was estimated by the rrBLUP model for resistance to the three SCN HG Types, 0, 2.5.7, and 1.2.3.5.6.7 (**Supplementary Table 13**, **Supplementary Figure 14**). As we did for GP estimations, the GH was estimated, using six different ratios of the training set: the testing set 19:1, 9:1, 4:1, 7:3, and 1:1, as 5, 10, 20, 30, 40, and 50% of the testing set in the GWAS panel with (1) all 4,654 SNPs (top in both **Supplementary Table 13**, **Supplementary Figure 14**), (2) 20 GWAS-derived SNP markers (middle), and (3) nine different SNP number sets from 20 SNPs to all 4,654 SNPs (bottom) in cross-prediction. The averaged GH was 22.4, 12.1, and 5.4% for three HG Types, respectively, in all 4,654 SNPs; 28.1, 22.1, and 6.1% in 20 SNP markers; and 13.7, 10.5, and 3.2% in the nine different SNP number sets from 20 SNPs to all 4,654 SNPs in cross-prediction. The results showed that GH was highest for resistance to HG Type 0, middle for HG Type 2.5.7, and lowest for HG Type 1.2.3.5.6.7, and the GWAS-derived 20 SNP marker set had higher GH (**Supplementary Table 13**).

Genetic Diversity Analysis for the SCN-Resistant Germplasm Accessions

There were 47 resistant accessions with FI < 20.0 for resistance to HG Type 0 (**Supplementary Table 1**). Among the 47 accessions, 10 had FI > 30.0 for resistance to HG Type 2.5.7, although they had FI values < 20.0 for resistance to both HG Type 0 and 1.2.3.5.6.7. These 10 accessions were not recognized as broadly resistant lines and were dropped from further genetic diversity analysis. Among the 37 accessions, 27 accessions were originally collected from Mexico, two from Colombia, three from Costa Rica, two from Ecuador, one from Guatemala, and two from Peru (**Table 1**) indicating that the SCN resistance was mainly distributed among Mesoamerican accessions in this study.

The 37 accessions formed two clusters as I and II (**Figure 8**, **Table 1**). Group I consisted of 27 accessions, which were mainly from Mexico, plus two from Colombia, one from Costa Rica,

and one from Guatemala. All of the 27 accessions also belonged to Cluster Q1, based on the population structure and genetic analyses of all 315 accessions. Group II had 10 accessions, including four from Mexico, two from Costa Rica, two from Ecuador, and two from Peru (**Figure 8**, **Table 1**). Among the 10 accessions in II, seven belonged to Q2, with a membership coefficient > 70%, and three to Q1 or Q1(2) with a membership coefficient of 30%. The latter three accessions, PI430204, PI346960, and PI345576, had Q1 membership coefficients > 66% based on the population structure and genetic analysis in all 315 accessions. In this phylogenetic tree of the 37 accessions (**Figure 8**), the three accessions, PI430204, PI346960, and PI345576, were clustered to group II but diverged from the cluster. The three accessions plus PI417657 more likely belonged to a subpopulation between clusters I and II, indicating the four accessions combined genetic backgrounds of both clusters (I and II) and the two subpopulations of common bean based on STRUCTURE 2.3.4 analysis.

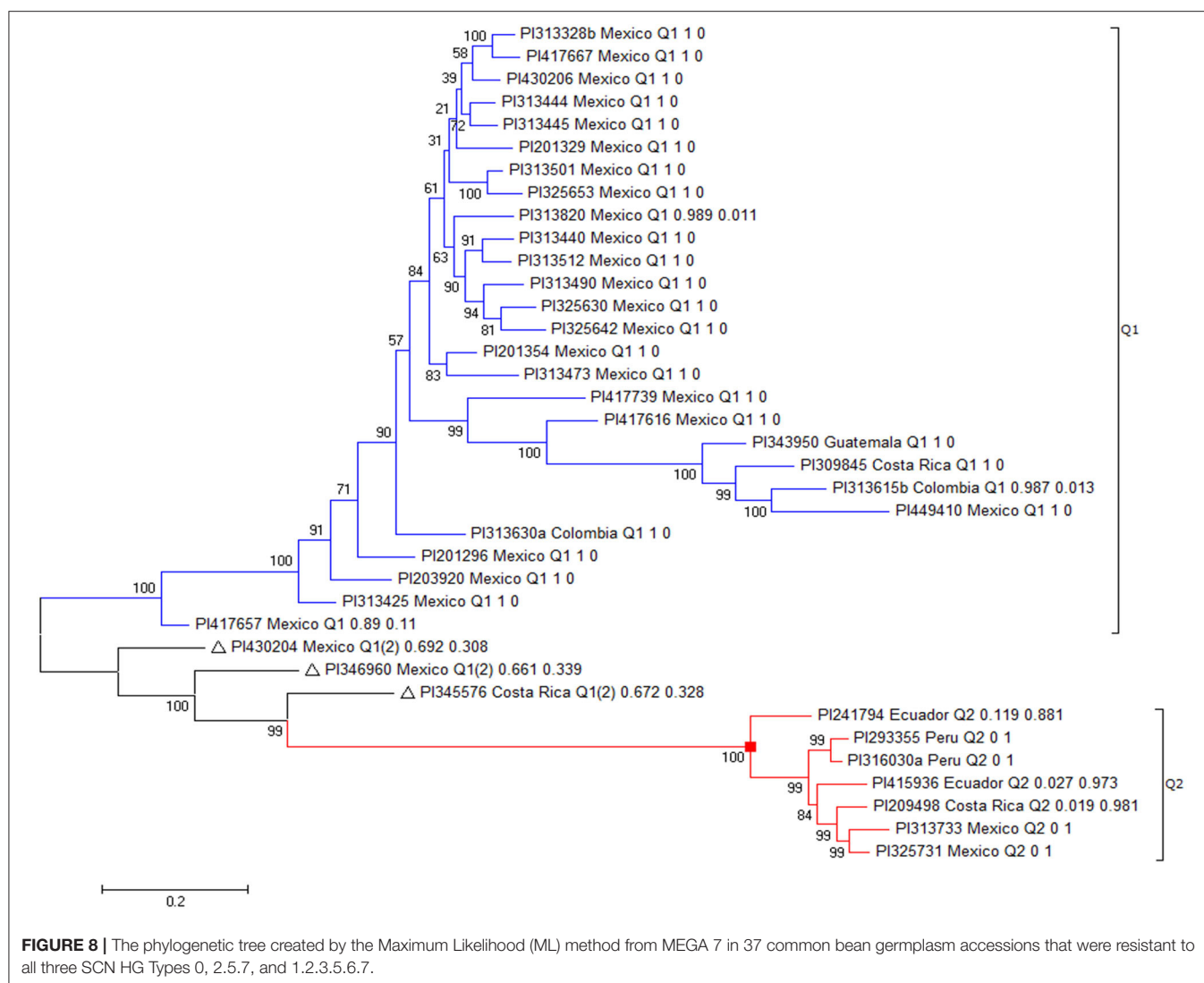
DISCUSSION

Genetic Diversity and Population Structure

In this study, the common bean population structure was examined among 315 common bean germplasm accessions belonging to the USDA *P. vulgaris* core collection, using the Markov Chain Monte Carlo iterations in STRUCTURE 2.3.4. The 315 accessions can be divided into two larger populations (Q1 and Q2 clusters) or into three subpopulations (G1 to G3 plus admixture) (**Figure 2**, **Supplementary Figure 3**, **Supplementary Table 1**).

Based on the two broader populations (Q1 and Q2) in the core collection, Q1 was the larger cluster with 77% (241/315) of accessions and consisted of germplasm mainly from Mexico (145), Guatemala (25), Colombia (20), Costa Rica (13), Nicaragua (12), El Salvador (11), and Honduras (7), with 60, 10, 8, 5, 5, and 3%, respectively (**Supplementary Table 14**). Q2 consisted of germplasm mainly from Mexico (15), Colombia (15), Peru (14), and Ecuador (11), with 22, 22, 21, and 16%, respectively (**Supplementary Table 14**). Most of the germplasm accessions from Central America, including Nicaragua (92.3%), Mexico (89.0%), Guatemala (83.3%), El Salvador (84.6%), Costa Rica (76.5%), and Honduras (77.8%) belonged to Q1; most accessions from South America, including Bolivia (only one accession), Peru (82.4%), and Ecuador (68.8%) belonged to Q2; and Colombia accessions belonged to both Q1 and Q2 with 57.1% to Q1 and 42.9% to Q2 (**Supplementary Table 13**).

Common bean consists of two geographic, diverged gene pools, namely the Andean and Middle American pools (Gepts and Bliss, 1985; Gepts et al., 1986, 2019; Koenig and Gepts, 1989; Koinange and Gepts, 1992; Beebe et al., 1997, 2000; Blair et al., 2009, 2012; Kwak and Gepts, 2009; Bitocchi et al., 2012, 2013; McClean et al., 2012; Schmutz et al., 2014; Campa et al., 2018; Kuzay et al., 2020). The analysis confirmed the presence of two populations (two clusters) among these 315 accessions but notes that the germplasm accessions from Nicaragua, Mexico, Guatemala, El Salvador, Costa Rica, Honduras, Colombia, Ecuador, and Peru include the members



of both clusters (populations), indicating that both gene pools existed in these countries.

Based on the three clusters (populations G1 to G3) in the 315 accessions, G1 had accessions mainly from Mexico (32 accessions), Colombia (13), Costa Rica (12), Nicaragua (11), and El Salvador (8), with 33, 13, 12, 11, 10, and 8%, respectively (**Supplementary Table 13**). G2 consisted of the accessions mainly from Mexico (111) and Guatemala (10), with 80 and 7%, respectively (**Supplementary Table 13**). G3 came from Mexico (15), Colombia (15), Peru (13), and Ecuador (11), with 23, 22, 20, and 17%, respectively (**Supplementary Table 13**). Besides, 14 accessions (4%) of the panel were admixed (**Supplementary Table 13**). For each country, most of the germplasm accessions from the United States (only one accession), Nicaragua (84.6%), Costa Rica (70.6%), El Salvador (61.5%), and Honduras (55.6%) belonged to G1. Most Mexico accessions (68%) belonged to G2; and most accessions from Peru (77%), Ecuador (69%), and Bolivia (only one accession) belonged to G3. The accessions from Guatemala and Colombia belonged

to three populations (Q1, Q2, and Q3); 23% of accessions from Guatemala were admixed (**Supplementary Table 13**). The three Q populations matched those in the report by Kuzay et al. (2020) when $K = 3$ (**Supplementary Table 1**). Furthermore, nearly half of the resistant accessions in this core collection belonged to the Middle American gene pool and the Durango ecogeographic race within this gene pool, although some resistant accessions were also identified in race Mesoamerica of the Middle American gene pool and races Nueva Granada and Peru of the Andean gene pool (**Supplementary Table 15**). Based on these results, we used the three Q-matrices for GWAS in all set of the 315 accessions to identify SNP markers associated with SCN resistance in this study.

Genome-Wide Association Study and SNP Marker Identification for SCN Resistance

In this GWAS study, six, 11, and four SNPs were identified to be associated with resistance to HG Types, 0, 2.5.7, and 1.2.3.5.6.7, respectively (**Table 2**). The six SNPs for HG Type 0 resistance

were newly identified markers for resistance to HG Type 0 (race 6) based on their location on chromosomes (**Table 2**). However, in the region of the two markers, ss715647158 and ss715649511 on Pv07, Jain et al. (2019) also reported an SNP marker ss715648793 (**Supplementary Table 16**) in the region, further validating a QTL in this region for HG Type 0 resistance. The SNP marker, ss715647549, was significantly associated with HG Type 0 resistance in two association panels, all.set, and Q1 (**Table 2**), and Jain et al. (2019) also reported six SNPs nearby (**Supplementary Table 16**), suggesting that there is a QTL on Pv11 for resistance to HG Type 0. Near ss715640464 on Pv04 (distance of 8.98 Kbp), a gene model Phvul.004G104700 of the disease resistance family protein/LRR family protein was found (**Table 3**) in the same LD region (**Figure 8**), suggesting that Phvul.004G104700 may be associated with the HG Type 0 resistance, but this observation needs to be validated.

For the 11 SNPs with resistance to HG Type 2.5.7, three are located on Pv01, two on Pv02, one on Pv03, one on Pv07, three on Pv09, and one on Pv11 (**Table 2**). The 11 SNPs are newly identified markers for resistance to HG Type 2.5.7. However, at the ss715639285 region on Pv02 and the ss715645642 region on Pv09, Jain et al. (2019) reported associated SNP markers for HG Type 0 resistance (**Supplementary Table 16**) and Wen et al. (2019) reported a close SNP marker on Pv09 for resistance to HG Type 2.5.7 resistance (**Supplementary Table 16**), suggesting that there are QTLs in the regions for SCN resistance. Further studies are needed to validate the broad resistance to multiple HG Types associated with these SNP markers.

The four SNPs with resistance to HG Type 1.2.3.5.6.7 were located on Pv03, 06, 10, and 11 (**Table 2**), and they are newly identified in this study. However, close to the ss715647109 region on Pv06, Jain et al. (2019) reported an SNP marker, ss715645673, associated with HG Type 0 resistance (**Supplementary Table 16**), indicating there may be a QTL in the region, but whether the QTL is associated with resistance to the two different HG Types needs to be further validated. Another SNP, ss715639563 at 46,491,205 bp on Pv11 for HG Type 1.2.3.5.6.7 resistance (**Table 2**), was close (distance ~1.84 Mbp) to ss715647549 at 44,651,807 bp, suggesting a QTL existed in the region, but whether this QTL is associated with resistance to both HG Types needs further study. However, based on the LD analysis (**Supplementary Figure 8F**, bottom right), the two SNPs, ss715647109 and ss715647549, are located in two different LD regions, suggesting that there are different genes or alleles for resistance to HG Type 0 and 2.5.7.

One SNP, ss715639339 at 12,175,377 bp on Pv09 was associated with both HG Type 0 and HG Type 2.5.7 resistance in two association panels, all.set, and Q1 (**Table 2**). Another SNP, ss715640389, at 12,154,448 bp on Pv09 was associated with resistance to HG Type 2.5.7 (**Table 2**). The two SNPs are very close to each other (within 20.929 Kbp) and located in the same LD region (**Supplementary Figure 8C**), suggesting that a QTL exists for SCN resistance, but further studies are needed to determine whether the QTL is associated with resistance to the two different HG Types.

So far, there are only two GWAS reports for SCN resistance in common beans (Jain et al., 2019; Wen et al., 2019). Wen et al. (2019) conducted GWAS in 363 accessions of the USDA

common bean core set for resistance to SCN HG types 2.5.7 and 1.2.3.5.6.7, using 84,416 SNPs obtained with GBS. They found five SNPs on Pv01 and one on Pv09, associated with resistance to HG Type 2.5.7, and only one SNP on Pv07, associated with HG Type 1.2.3.5.6.7 resistance. The five SNP markers with resistance to HG Type 2.5.7 were located at 10,061,925 bp, 18,388,378 bp, 18,388,392 bp, 18,388,403 bp, and 18,388,408 bp on Pv01 of the *P. vulgaris* G19833 Pvulgaris v1.0 reference sequence (Schmutz et al., 2014), with *P*-value from 1.02×10^{-6} to 4.94×10^{-6} , and another one on Pv09 at 35,068,146 bp with *P*-value 1.80×10^{-6} . The SNP marker for resistance to HG Type 1.2.3.5.6.7 was located at 44,761,605 bp on Pv07. We used the SCN phenotypic data from the report by Wen et al. (2019), but a different set of SNPs in BARCBean6K_3 BeadChips (Song et al., 2015) to redo the GWAS analysis. We did not identify the same SNP markers but identified the SNP markers in the same regions for resistance to SCN HG Type 2.5.7, but not for HG Type 1.2.3.5.6.7 resistance (**Table 2**, **Supplementary Tables 6, 7**). The two SNPs, ss715640034 and ss715639810, located at 18,874,808 bp and 20,450,707 bp on Pv01 and two SNPs, ss715645642 and SS71549401 located at 33,052,539 bp and 33,956,905 bp on Pv09 (**Table 6**) were located in similar regions reported by Wen et al. (2019), suggesting that there are QTLs for HG Type 2.5.7 resistance in these regions.

Jain et al. (2019) conducted GWAS in 317 accessions of USDA common bean core collection with SCN HG Type 0 and found 14 significant SNP markers on Pv 04, 05, 06, 07, 08, 10, and 11 in the Middle American subpopulation (179 accessions) and 23 SNP markers on Pv 01, 02, 07, 08, 09, and 11 for the Andean subpopulation (138 accessions). However, we could not find any of the 37 SNPs with LOD values greater than the significant threshold values of 4.97 in all.set, 4.84 in Q1, and 4.52 in Q2 for the resistance to HG Type 0, 2.5.7, or 1.2.3.5.6.7, respectively (**Supplementary Table 16**). Nevertheless, 11 of the 37 SNPs had at least one LOD value >3.0 from GAPIT3 or TASSEL 5 and, also, a LOD score > 3.0 in *t*-tests for resistance to the population of HG Type 0 (**Supplementary Table 17a**). We did not retain them as significant associated SNP markers because each of the 11 SNPs did not have LOD values greater than the significant threshold, even <3.0 in any MLM model, although they may be associated with the resistance to HG Type 0 with a minor effect (**Supplementary Table 17a**). In addition, we observed nine and 10 SNPs with LOD values >3.0 in one or more models and *t*-tests as well (**Supplementary Tables 17b,c**), suggesting these SNPs have minor effects for resistance to either HG Type 2.5.7 or 1.2.3.5.6.7.

Candidate Gene Model

Wen et al. (2019) reported three gene models, PHAVU_001G248000g (amino acid transporter), PHAVU_001G247900g (α -SNAP protein), and PHAVU_001G247700g (wound inducible protein 12), located at 50,653,407–50,655,828 bp, 50,646,068–50,650,097 bp, and 50,629,261–50,630,123 bp respectively, on Pv01 of common beans to be associated with resistance to HG Type 2.5.7, which corresponded to three gene models in the *rhg1* region of soybean chr18 with 91%, 94%, and 88% identities. However, Wen et al. (2019) did not report any associated SNP marker in a 50 Mbp

region of chromosome Pv01; the closest gene model was located at 18,388,408 bp, which was 32 Mbp distance away from the three genes. The data of resistance to SCN HG Type 2.5.7 from Wen et al. (2019) did not confirm the *rhg1* in soybean existed in common beans for their study. Jain et al. (2019) also reported several candidate genes on Pvulgaris v1.0 Pv01 and Pv08, which had high similarity to the three genes of *rhg1* of soybean for SCN resistance, but they did not report any significant SNP marker located in the candidate gene regions, which were associated with the resistance to HG Type 0. Thus, their study could not confirm either that there is *rhg1* or *Rhg4* resistance in common beans. From the study, an SNP marker, ss715645939, was associated with HG Type 2.5.7, which was located at 48,772,176 bp on Pvulgaris v1.0 Pv01, at a distance of around 1.9 Mbp from the three *rhg1* paralog genes in common beans (Supplementary Table 6). The low LOD values of the SNP marker (LOD < 4 in all six MLM models and 5.0 in GLM and 5.12 in SMR, Supplementary Table 6) cast doubt about resistance to SCN HG Types at this location.

From this study, two LRR gene models, Phvul.004G099300 and Phvul.010G018300 were identified as candidates for SCN resistance. Phvul.004G099300 (disease resistance family protein/LRR family protein) at 33,316,658–33,320,257 bp on Pv04 was associated with HG Type 0 resistance, and Phvul.010G018300 (LRR protein kinase family protein) at 2,832,211–2,839,756 on Pv10 was associated with resistance to HG Type 1.2.3.5.6.7 (Table 3). However, the LRR gene in the *rhg1* region on chr 18 in soybean was not involved in SCN resistance (Mitchum, 2016). Further studies are needed to validate whether the two genes are responsible for the SCN resistance in common beans.

Genomic Prediction

Genomic prediction accuracy, using the Pearson's correlation coefficient (r) between the GEBV and the observed values, has been the main parameter to measure the performance of GS (Jarquin et al., 2014, 2016; Zhang J. P. et al., 2016; Qin et al., 2019; Ravelombola et al., 2019, 2020, 2021; Wen et al., 2019; Ali et al., 2020; Keller et al., 2020). The PA is affected by several factors, such as the trait itself with its heritability, marker number, and the marker associated with the trait, and is also affected by GS models, marker density, the level of LD, QTL number, the population size, and the relationship between training population and testing population (Jarquin et al., 2016; Ali et al., 2020; Keller et al., 2020). In this study, five scenarios were tested for genomic PA: (1) different ratios of the training set and the testing set (validation set), (2) different SNP numbers, (3) three association panels, (4) the use of GWAS-derived significant SNP markers, and (5) different GP models for resistance to three SCN HG Types.

In this study, GP was performed, using six different ratios of the training set: the testing set 19:1, 9:1, 4:1, 7:3, and 1:1, as 5, 10, 20, 30, 40, and 50% of the testing set in the panel. The six tests showed similar PA (averaged r -values). A small decrease of the r -value was observed in most cases with testing sets of 40% or higher. But the 5% "Testing set" (19:1 in the training set: the testing set) had the largest variance, and 50% had the

smallest. The averaged r -values decreased from 5 to 50% (Table 4, Supplementary Tables 9, 11, and Supplementary Figures 6, 9). The study showed that 10, 20, and 30% of the testing set size (as the same 9:1, 4:1, and 7:3 of the training set: the testing set) are good to be used in GS for HG Type resistance in common beans. Keller et al. (2020) reported that the training set of <30% could reduce PA due to an insufficiently sized training set that resulted in overfitting of the model; they also reported that a training set > 80% can lead to large variation between cross-validations due to an excessively small validation set. The results showed similar trends but 10% of the testing set size (i.e., training set size = 90%) was acceptable to GS. Ravelombola et al. (2021) reported that the average GS accuracy was similarly based on the r -values at 2-fold [training set: testing set (validation set) = 1:1], 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, and 8-fold cross-validation for growth habit, flowering time, and a grain yield in a multi-parent advanced generation intercross (MAGIC) cowpea population under drought condition, but a slightly higher averaged r -value was observed in 7-fold cross-validation for 100-seed weight, perhaps associated with the higher heritability of seed weight (Nienhuis and Singh, 1988).

In this study, GP was also performed with nine different SNP number sets from 20 to all 4,654 SNPs in cross-prediction for resistance to three HG Types, using five GP models: rrBLUP, Bayes A, Bayes B, BL, and BRR (Table 5). PA decreased when 100 or less SNPs were used for HG Type 0 resistance and when 200 or less SNPs were used for HG Type 2.5.7 resistance, but PA did not decrease for HG Type 1.2.3.5.6.7 resistance (Table 5, Figure 7, Supplementary Table 10, Supplementary Figure 10). Overall, the results suggest that > 200 SNPs should be used for GS. Wen et al. (2019) reported the average PA estimated by cross-validation was 0.52 and 0.41 for SCN HG Type 2.5.7 and HG Type 1.2.3.5.6.7, respectively, when 5,000 SNPs or more were used and showed a decrease when 1,000 SNPs were used. In most of the reports, the smaller the number of SNPs used, the lower the PA was (Jarquin et al., 2014, 2016; Zhang J. P. et al., 2016; Wen et al., 2019; Ali et al., 2020). Zhang J. P. et al. (2016) estimated PA (r -value) of seed size based on 309 soybean accessions and reported $r = 0.85$ when 2,000 SNPs or 31,045 SNPs were included; $r = 0.8$ when 1,000 SNPs or 500 SNPs were used.

In this study, using GWAS-derived SNP markers led to the highest GP accuracy for resistance to all three SCN HG Types (Supplementary Tables 9, 12, Supplementary Figure 12). Ali et al. (2020) estimated the prediction accuracy of various GS models on yield and yield-related traits in wheat; they reported that the GWAS-derived markers improved PA in most cases. Zhang J. P. et al. (2016) conducted GWAS and identified 48 SNPs on 12 chromosomes associated with soybean seed size. Based on GWAS, they reported that the r -values ranged from 0.64 to 0.74 when 5, 10, and 15 of the 48 SNP markers were used, which were 25% higher than those calculated from the same number of randomly selected SNPs. Qin et al. (2019) reported that the average correlation coefficient (r) among 15 amino acids between the observed values (each amino acid content) and the GEBVs predicted ranged from 0.18 to 0.61 when all 23,279 SNPs were used, from 0.45 to 0.68 when 231 SNP markers, associated with one or more amino acid from GWAS were used; and 0.33 to 0.54

when only the associated SNP markers with the specific amino acid content were used, using RR-BLUP in rrBLUP software. Spindel et al. (2016) developed a GS model (GS + *de novo* GWAS) that combines RR-BLUP with GWAS-derived markers, which were fitted as fixed effects on the RR-BLUP training data and found that this new model outperformed other models, RR-BLUP, Bayesian LASSO (BL), Reproducing Kernel Hilbert Spaces (RKHS) and RF, and multiple linear regression (MLR) for a variety of traits in multiple environments. Thus, using GWAS-derived SNP markers to perform GS is an approach combining MAS and GS that can be used in the real-world breeding program, although the predictive ability may be biased, using SNP markers from GWAS to predict the GEBVs in the same GWAS panel. The real GP will be lower if conducting predictions in other panels with different individuals. We have tested many traits in several crops and find it is a practical approach to do genome breeding, using GWAS-derived SNP markers (Qin et al., 2019; Ravelombola et al., 2019, 2020, 2021). Therefore, an approach combining MAS and GS through GEBVs, using associated SNP markers (Spindel et al., 2016; Zhang J. P. et al., 2016; Qin et al., 2019; Ravelombola et al., 2019, 2020, 2021; Ali et al., 2020) will be a good choice to do molecular breeding for SCN resistance in common beans and, also, for other quantitative traits in other plant species.

In addition, GA is affected by the trait self, such as heritability. The GH has been estimated and reported in animals and plants such as heifers (Nawaz et al., 2018), soybean (Xavier and Rainey, 2020), and safflower (Zhao et al., 2021). de los Campos et al. (2015) developed whole-genome regression methods to estimate the GH, which was defined as the proportion of variance of a trait that can be explained (in the population) by linear regression on a set of markers. In this study, the GH was also estimated by the rrBLUP model for resistance to the three SCN HG Types, 0, 2.5.7, and 1.2.3.5.6.7 (**Supplementary Table 13**, **Supplementary Figure 14**), as we did for GP estimations. The results indicated that the higher GH, the higher GP, similar as reported by Xavier and Rainey (2020) for yield and related traits in soybeans.

Utility of Common Bean Resistance Accessions

From this study, 15 out of 315 (4.8%) common bean accessions were resistant to SCN, with FI ranging from 4.8 to 10; 62 (19.7%) accessions were moderately resistant ($10 < \text{FI} < 30$) for HG Type 0 (race 6). The 15 resistant accessions were PI343950, PI313630, PI313328, PI201329, PI201354, PI313445, PI313440, PI313444, PI319684, PI417616, PI313501, PI325614, PI430206, PI313733, and PI269209, which will be preferred sources for resistance to HG Type 0 (race 6).

To select common bean accessions with resistance to multiple SCN HG Types, we combined the data of the SCN resistance to HG Types, 2.5.7 and 1.2.3.5.6.7 from the Wen et al. (2019) report and the data. We then selected 37 accessions, having broad resistance with $\text{FI} < 20$ to both HG Types, 0 and 1.2.3.5.6.7, and $\text{FI} < 30$ to HG Type 2.5.7 (**Table 1**). The genetic diversity of the 37 accessions showed similar to the genetic organization of the

entire 315 accession collections (**Figure 8**, **Table 1**). Most of the resistant accessions belonged to the ecogeographic race Durango of the Middle American gene pool, although other gene pools or races also contained SCN resistance. The accessions with the highest resistance to multiple HG Types (with $\text{FI} < 12$ to the three HG Types) were PI201329, PI201354, PI313445, PI325642 (all race Durango), PI313733 (Andean admixed), and PI417616 (admixed) (**Table 1**, **Supplementary Table 1**).

These resistant accessions can be used in common bean breeding programs as parents to develop new cultivars with resistance to multiple SCN HG Types. In this study, we observed that the SCN resistance commonly existed in common bean accessions. There were 15 out of 315 (4.8%) common bean accessions resistant to HG Type 0 (race 6) with $\text{FI} < 10$ (**Supplementary Table 1**). Based on the report by Wen et al. (2019), 19 out of 363 accessions (5.2%) were resistant to HG Type 2.5.7, and 160 out of 363 (44.1%) resistant to HG Type 1.2.3.5.6.7 with $\text{FI} < 10$.

Interestingly, there were much more common bean lines resistant to HG Type 1.2.3.5.6.7 than HG Type 2.5.7 and HG Type 0. This contrasts to the SCN resistance in soybean, which has fewer lines resistant to HG 1.2.3.5.6.7 as compared with HG Type 2.5.7, and much fewer lines as compared with HG Type 0 because a population of HG Type 1.2.3.5.6.7 generally has broader virulence than a population of HG Type 2.5.7 or HG Type 0. Although the FI on the HG Type indicator lines of the two SCN populations used by Wen et al. (2019) was not reported, it is possible that the mechanisms of SCN resistance differed between soybeans and common beans. If this is true, the different and broad-spectrum SCN resistance in common beans potentially provides excellent sources of SCN resistance to soybeans. SCN has been the most damaging pest in soybeans. Only a few sources available for resistance to multiple HG Types, particularly for resistance to HG Type 2.5.7 and 1.2.3.5.6.7, but none of them has been successfully deployed in commercial soybean cultivars. After the discovery of the SCN resistance genes in common beans, it will be possible to transfer the genes from common beans to soybeans through a transgenic approach.

CONCLUSION

In this study, 15 accessions of the USDA common bean core collection were observed for the resistance to SCN HG Type 0 with FI at 4.8 to 9.4; six SNP markers, located on chromosomes Pv 04, 06, 07, 07, 09, and 11, respectively, were significantly associated with the resistance to this SCN HG Type 0. GWAS was also conducted for resistance to HG Type 2.5.7 and HG Type 1.2.3.5.6.7 based on published phenotypic data and the genotypic data from the BARCBean6K_3 chip. Eleven SNPs were associated with HG Type 2.5.7 resistance on chromosomes Pv01, 02, 03, 07, 09, and 11, and four SNPs with HG Type 1.2.3.5.6.7 resistance on chromosomes Pv 03, 06, 10, and 11. A gene model of the disease resistance family protein/LRR protein family, Phvul.004G104700, was close to the SNP marker ss715640464 at a distance of 8.98 Kbp in the same LD region of chromosome Pv04, suggesting that Phvul.004G104700 may be a candidate gene for the HG Type 0

resistance. GP was performed for resistance to three HG Types, using eight GP models (rrBLUP, gBLUP, Bayes A, Bayes B, BL, BRR, RF, and SVM), with BL showing the most promising results in terms of PA. The results showed that 400 SNPs or more had similar GS efficiency for resistance to either HG Type 0, 2.5.7, or 1.2.3.5.6.7, and the set of 20 significant SNP markers had the highest PA for GP. The six sets of different testing set sizes from 5 to 50% had similar *r*-values. Except for gBLUP (lower PA), all other seven models had similar PA. The averaged *r*-values were 0.40 for HG Type 0 resistance, 0.34 for HG Type 2.5.7, and 0.32 for HG Type 1.2.3.5.6.7. This study provides basic information for breeders to develop SCN-resistant common bean cultivars, using the USDA core germplasm accessions through MAS and GS in common beans.

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

SC was the principal investigator (PI), performed phenotypic data collection and analysis, and revised and wrote part of the manuscript. TM and AS were the Co-PI of the project. QS developed the SNP Chip and genotyped the accessions. PG performed genotypic data collection, using the SNP Chip. AS

performed genomic and statistical analysis and wrote the draft of the manuscript. HX assisted in the data analysis. All the authors have edited, reviewed, and approved the manuscript.

FUNDING

This research was supported by the Minnesota Department of Agriculture AGRI (Agricultural Growth, Research, and Innovation) Program and USDA Germplasm Evaluation grant with Agreement Number/FAIN: 58-2090-8-053 and Project No. 2090-21000-032-11S; Accessions No. 434981.

ACKNOWLEDGMENTS

The authors would like to thank Wayne Gottschalk, Cathryn Johnson, Hannah Neigebauer, Yen N. Vuong, Fengyu Shi, Jeffery Ballman, and undergraduate students for their assistance in conducting the phenotyping greenhouse experiments and data collection. Thanks to Zhiwu Zhang and Jiabo Wang who provided the GAPIT3 tool for us to run GWAS and GP analysis and to Waltram Ravelombola for his writing the R-codes for us to run the GP analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.624156/full#supplementary-material>

REFERENCES

- Albrecht, T., Wimmer, V., Auinger, H. J., Erbe, M., Knaak, C., Ouzunova, M., et al. (2011). Genome-based prediction of testcross values in maize. *Theor. Appl. Genet.* 123, 339–350. doi: 10.1007/s00122-011-1587-7
- Ali, M., Zhangm, Y., Rasheed, A., Wang, J., and Zhang, L. (2020). Genomic prediction for grain yield and yield-related traits in Chinese winter wheat. *Int. J. Mol. Sci.* 21:1342. doi: 10.3390/ijms21041342
- Allen, T. W., Bradley, C. A., Sisson, J. A., Byamukama, E., Chilvers, M. I., Coker, C. M., et al. (2017). Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. *Plant Health Prog.* 18, 19–27. doi: 10.1094/PHP-RS-16-0066
- Bao, Y., Vuong, T., Meinhardt, C., Tiffin, P., Denny, R., Chen, S., et al. (2014). Potential of association mapping and genomic selection to explore PI 88788 derived soybean cyst nematode resistance. *Plant Genome* 7, 1–13. doi: 10.3835/plantgenome2013.11.0039
- Barili, L. D., do Vale, N. M., Silva, F. F., de Souza Carneiro, J. E., de Oliveira, H. R., Vianello, R. P., et al. (2018). Genome prediction accuracy of common bean via Bayesian models. *Ciência Rural Santa Maria* 48:e20170497. doi: 10.1590/0103-8478cr20170497
- Barrett, J. C., Fry, B., Maller, J., and Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21, 263–265. doi: 10.1093/bioinformatics/bth457
- Beebe, S., Skroch, P. W., Tohme, J., Duque, M. C., Pedraza, F., and Nienhuis, J. (2000). Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Sci.* 40, 264–273. doi: 10.2135/cropsci2000.401264x
- Beebe, S., Toro, O., González, A., Chacón, M., and Debouck, D. (1997). Wild-weed-complexes of common bean (*Phaseolus vulgaris* L., Fabaceae) in the Andes of Peru and Colombia, and their implications for conservation and breeding. *Genet Resour Crop Evol.* 44, 73–91. doi: 10.1023/A:1008621632680
- Bernardo, R., and Yu, J. M. (2007). Prospects for genomewide selection for quantitative traits in maize. *Crop Sci.* 47, 1082–1090. doi: 10.2135/cropsci2006.11.0690
- Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodriguez, M., Biagetti, E., et al. (2013). Molecular analysis of the parallel domestication of the common bean (*Phaseolus vulgaris*) in Mesoamerica and the Andes. *New Phytol.* 197, 300–313. doi: 10.1111/j.1469-8137.2012.04377.x
- Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., et al. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proc. Natl. Acad. Sci. U.S.A.* 109, E788–E796. doi: 10.1073/pnas.1108973109
- Blair, M. W., Cortes, A. J., Penmetza, R., Farmer, A., Carrasquilla-Garcia, N., and Cook, D. R. (2013). A high-throughput SNP marker system for parental polymorphism screening, and diversity analysis in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 126, 535–548. doi: 10.1007/s00122-012-1999-z
- Blair, M. W., Díaz, L. M., Buendía, H. F., and Myriam C Duque, M. C. (2009). Genetic diversity, seed size associations and population structure of a core collection of common beans (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 119, 955–972. doi: 10.1007/s00122-009-1064-8
- Blair, M. W., and Izquierdo, P. (2012). Use of the advanced backcross-QTL method to transfer seed mineral accumulation nutrition traits from wild to Andean cultivated common beans. *Theor. Appl. Genet.* 125, 1015–1031. doi: 10.1007/s00122-012-1891-x
- Blair, M. W., Soler, A., and Cortés, A. J. (2012). Diversification and population structure in common beans (*Phaseolus vulgaris* L.). *PLoS ONE* 7:e49488. doi: 10.1371/journal.pone.0049488

- 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
- Campa, A., Murube, E., and Ferreira, J. J. (2018). Genetic diversity, population structure, and linkage disequilibrium in a Spanish common bean diversity panel revealed through genotyping-by-sequencing. *Genes* 9:518. doi: 10.3390/genes9110518
- Castro, S. A. D., Gonçalves-Vidigal, M. C., Gilio, T. A. S., Lacanallo, G. F., Valentini, G., Martins, V. D. R., et al. (2017). Genetics and mapping of a new anthracnose resistance locus in Andean common bean Paloma. *BMC Genomics* 18:306. doi: 10.1186/s12864-017-3685-7
- Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142, 169–196. doi: 10.1007/s10681-005-1681-5
- Collard, B. C. Y., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. B.* 363, 557–572. doi: 10.1098/rstb.2007.2170
- Cortés, A. J., Chavarro, M. C., and Blair, M. W. (2011). SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 123, 827–845. doi: 10.1007/s00122-011-1630-8
- Cortés, A. J., Felipe, L. H., and Daniela, O. R. (2020). Predicting thermal adaptation by looking into populations' genomic past. *Front. Genet.* 11:564515. doi: 10.3389/fgene.2020.564515
- de los Campos, G., Sorensen, D., and Gianola, D. (2015). Genomic heritability: what is it? *PLoS Genet.* 11:e1005048. doi: 10.1371/journal.pgen.1005048
- Desta, Z. A., and Ortiz, R. (2014). Genomic selection: genome-wide prediction in plant improvement. *Trends Plant Sci.* 19, 592–601. doi: 10.1016/j.tplants.2014.05.006
- Duan, Y. X., Zheng, Y. N., Chen, L. J., Zhou, X. M., Wang, Y. Y., and Sun, J. N. S. (2009). Effects of abiotic environmental factors on soybean cyst nematode. *Agric. Sci. China* 8, 317–325. doi: 10.1016/S1671-2927(08)60215-1
- 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
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R Package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024
- 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
- Faghihi, J., and Ferris, J. M. (2000). An efficient new device to release eggs from *Heterodera glycines*. *J. Nematol.* 32, 411–413.
- Galeano, C. H., Cortés, A. J., Fernández, A. C., Soler, A., Franco-Herrera, N., Makunde, G., et al. (2012). Gene-based single nucleotide polymorphism markers for genetic and association mapping in common bean. *BMC Genetics* 13:48. doi: 10.1186/1471-2156-13-48
- Garzón, L. N., Ligarreto, G. A., and Blair, M. W. (2008). Molecular marker-assisted backcrossing of anthracnose resistance into andean climbing beans (*Phaseolus vulgaris* L.). *Crop Sci.* 48, 562–570. doi: 10.2135/cropsci2007.08.0462
- Gepts, P., and Bliss, F. A. (1985). F1 hybrid weakness in the common bean: Differential geographic origin suggests two gene pools in cultivated bean germplasm. *J. Hered.* 76, 447–450.
- Gepts, P., Kuzay, S., and Hamilton, P. (2019). An Evaluation of the Representativity of the USDA Core Collection of Common Bean as Assessed by SNP Diversity, UC. Davis, CA: Dryad.
- Gepts, P., Osborn, T. C., Rashka, K., and Bliss, F. A. (1986). Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. *Econ. Bot.* 40, 451–468. doi: 10.1007/BF02859659
- Hagerty, C. H., Cuesta-Marcos, A., Cregan, P., Song, Q. J., McClean, P., and Myers, J. R. (2016). Mapping snap bean pod and color traits, in a dry bean x snap bean recombinant inbred population. *J. Am. Soc. Hortic. Sci.* 141, 131–138. doi: 10.21273/JASHS.141.2.131
- Hagerty, C. H., Cuesta-Marcos, A., Cregan, P. B., Song, Q., McClean, P., Noffsinger, S., et al. (2015). Mapping *Fusarium solani* and *Aphanomyces euteiches* root rot resistance and root architecture quantitative trait loci in common bean. *Crop Sci.* 55, 1969–1977. doi: 10.2135/cropsci2014.11.0805
- Hayes, B. J., Bowman, P. J., Chamberlain, A. C., Verbyla, K., and Goddard, M. E. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genet. Sel. Evol.* 41:51. doi: 10.1186/1297-9686-41-51
- Heffner, E. L., Jannink, J. L., Iwata, H., Souza, E., and Sorrells, M. E. (2011). Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* 51, 2597–2606. doi: 10.2135/cropsci2011.015.0253
- 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
- Holland, J. (2003). Estimating and interpreting heritability for plant breeding: an update. *Plant Breed. Rev.* 22:9–112. doi: 10.1002/9780470650202.ch2
- Hoyos-Villegas, V., Song, Q. J., and Kelly, J. D. (2017). Genome-wide association analysis for drought tolerance and associated traits in common bean. *Plant Genome* 10, 1–17. doi: 10.3835/plantgenome2015.12.0122
- Hoyos-Villegas, V., Song, Q. J., Wright, E. M., Beebe, S. E., and Kelly, J. D. (2016). Joint linkage QTL mapping for yield and agronomic traits in a composite map of three common bean RIL populations. *Crop Sci.* 56, 2546–2563. doi: 10.2135/cropsci2016.01.0063
- Huang, M., Liu, X. L., Zhou, Y., Summers, R. M., and Zhang, Z. W. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8:giy154. doi: 10.1093/gigascience/giy154
- Huang, X. H., Wei, X. H., Sang, T., Zhao, Q. A., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695
- Hurtado-Gonzales, O. P., Valentini, G., Gilio, T. A. S., Martins, A. M., Song, Q. J., and Pastor-Corrales, M. A. (2017). Fine mapping of Ur-3, a historically important rust resistance locus in common bean. *G3 Genes Genomes Genet.* 7, 557–569. doi: 10.1534/g3.116.036061
- Jain, S., Chittam, K., Brueggeman, R., Osorno, J. M., Richards, J., and Nelson, B. D. (2016). Comparative transcriptome analysis of resistant and susceptible common bean genotypes in response to soybean cyst nematode infection. *PLoS ONE* 11:e0159338. doi: 10.1371/journal.pone.0159338
- Jain, S., Poromarto, S., Osorno, J. M., McClean, P. E., and Nelson, B. D. (2019). Genome wide association study discovers genomic regions involved in resistance to soybean cyst nematode (*Heterodera glycines*) in common bean. *PLoS ONE* 14:e0212140. doi: 10.1371/journal.pone.0212140
- Jarquín, D., Kocak, K., Posadas, L., Hyma, K., Jedlicka, J., Graef, G., et al. (2014). Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC Genomics* 15:740. doi: 10.1186/1471-2164-15-740
- Jarquín, D., Specht, J., and Lorenz, A. (2016). Prospects of genomic prediction in the USDA soybean germplasm collection: historical data creates robust models for enhancing selection of accessions. *G3 Genes Genomes Genet.* 6, 2329–2341. doi: 10.1534/g3.116.031443
- Keller, B., Ariza-Suarez, D., de la Hoz, J., Aparicio, J. S., Portilla-Benavides, A. E., Buendia, H. F., et al. (2020). Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress. *Front. Plant Sci.* 11:1001. doi: 10.3389/fpls.2020.01001
- Koenig, R., and Gepts, P. (1989). Allozyme diversity in wild *Phaseolus vulgaris*: further evidence for two major centers of diversity. *Theoret. Appl. Genet.* 78, 809–817. doi: 10.1007/BF00266663
- Koenning, S. R., and Wraether, J. A. (2010). Suppression of soybean yield potential in the continental United States by plant diseases from 2006 to 2009. *Plant Health Prog.* 27, 1–6. doi: 10.1094/PHP-2010-1122-01-RS
- Koinange, E. M. K., and Gepts, P. (1992). Hybrid weakness in wild *Phaseolus vulgaris* L. *J. Heredity* 83, 135–139. doi: 10.1093/oxfordjournals.jhered.a111173
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Kuzay, S., Hamilton-Conaty, P., Palkovic, A., and Gepts, P. (2020). Is the USDA core collection of common bean representative of genetic diversity of the species, as assessed by SNP diversity? *Crop Sci.* 60, 1398–1414. doi: 10.1002/csc2.20032

- Kwak, M., and Gepts, P. (2009). Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). *Theor. Appl. Genet.* 118, 979–992. doi: 10.1007/s00122-008-0955-4
- Legarra, A., Robert-Granie, C., Croiseau, P., Guillaume, F., and Fritz, S. (2011). Improved Lasso for genomic selection. *Genet. Res.* 93, 77–87. doi: 10.1017/S0016672310000534
- Li, H., Peng, Z. Y., Yang, X. H., Wang, W. D., Fu, J. J., Wang, J. H., et al. (2013). Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* 45, U43–U72. doi: 10.1038/ng.2484
- Lin, Z., Hayes, B. J., and Daetwyler, H. D. (2014). Genomic selection in crops, trees and forages: a review. *Crop Pasture Sci.* 65, 1177–1191. doi: 10.1071/CP13363
- Lipka, A. E., Tian, F., Wang, Q. S., 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
- Liu, X. L., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. W. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12:e1005767. doi: 10.1371/journal.pgen.1005767
- López-Hernández, F., and Cortés, A. J. (2019). Last-generation genome-environment associations reveal the genetic basis of heat tolerance in common bean (*Phaseolus vulgaris* L.). *Front. Genet.* 10:954. doi: 10.3389/fgene.2019.00954
- Lv, J., Qi, J. J., Shi, Q. X., Shen, D., Zhang, S. P., Shao, G. J., et al. (2012). Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). *PLoS ONE* 7:e46919. doi: 10.1371/journal.pone.0046919
- Maenhout, S., De Baets, B., Haesaert, G., and Van Bockstaele, E. (2007). Support vector machine regression for the prediction of maize hybrid performance. *Theor. Appl. Genet.* 115, 1003–1013. doi: 10.1007/s00122-007-0627-9
- McClean, P. E., Terpstra, J., McConnell, M., White, C., Lee, R., and Mamidi, S. (2012). Population structure and genetic differentiation. Among the USDA common bean (*Phaseolus vulgaris* L.) core collection. *Genet. Resour. Crop Evol.* 59, 499–515. doi: 10.1007/s10722-011-9699-0
- Mitchum, M. G. (2016). Soybean resistance to the soybean cyst nematode *Heterodera glycines*: an update. *Phytopathology* 106, 1444–1450. doi: 10.1094/PHYTO-06-16-0227-RVW
- Moghaddam, S. M., Mamidi, S., Osorno, J. M., Lee, R., Brick, M., Kelly, J., et al. (2016). Genome-wide association study identifies candidate loci underlying agronomic traits in a Middle American diversity panel of common bean. *Plant Genome* 9, 1–21. doi: 10.3835/plantgenome2016.02.0012
- Moose, S. P., and Mumm, R. H. (2008). Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* 147, 969–977. doi: 10.1104/pp.108.118232
- 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
- Nawaz, M. Y., Jimenez-Krassel, F., Steibel, J. P., Lu, Y., Baktula, A., Vukasinovic, N., et al. (2018). Genomic heritability and genome-wide association analysis of anti-Müllerian hormone in Holstein dairy heifers. *J. Dairy Sci.* 101, 8063–8075. doi: 10.3168/jds.2018-14798
- Niblack, T. L., Arelli, P. R., Noel, G. R., Opperman, C. H., Orf, J. H., Schmitt, D. P., et al. (2002). A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *J. Nematol.* 34, 279–288.
- Nienhuis, J., and Singh, S. P. (1988). Genetics of seed yield and its components in common bean (*Phaseolus vulgaris* L.) of Middle American origin. II. Genetic variance, heritability, and expected response from selection. *Plant Breed.* 101, 155–163. doi: 10.1111/j.1439-0523.1988.tb00281.x
- Ogutu, J. O., Piepho, H. P., and Schulz-Streeck, T. (2011). A comparison of random forests, boosting and support vector machines for genomic selection. *BMC* 5(Suppl. 3):S11. doi: 10.1186/1753-6561-5-S3-S11
- Onogi, A., Watanabe, M., Mochizuki, T., Hayashi, T., Nakagawa, H., Hasegawa, T., et al. (2016). Toward integration of genomic selection with crop modelling: the development of an integrated approach to predicting rice heading dates. *Theor. Appl. Genet.* 129, 805–817. doi: 10.1007/s00122-016-2667-5
- Osorno, J. M., Vander Wal, A. J., Posch, J., Simons, K., Grafton, K. F., Pasche, J. S., et al. (2020). 'ND Falcon', a new pinto bean with combined resistance to rust and soybean cyst nematode. *J. Plant Regist.* 14, 117–125. doi: 10.1002/plr2.20025
- Poromarto, S. H., Nelson, B. D., and Goswami, R. S. (2010). Effect of soybean cyst nematode on growth of dry bean in the field. *Plant Dis.* 94, 1299–1304. doi: 10.1094/PDIS-05-10-0326
- Poromarto, S. H., Nelson, B. D., Goswami, R. S., and Welsh, M. (2012). Reproduction of soybean cyst nematode on accessions of the core collection of *Phaseolus vulgaris*. *Phytopathology* 102, 93–94.
- 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
- Qin, J., Shi, A., Song, Q., Li, S., Wang, F., Cao, Y., et al. (2019). Genome wide association study and genomic selection of amino acid contents in soybean seeds. *Front. Plant Sci.* 10:1445. doi: 10.3389/fpls.2019.01445
- Ravelombola, W., Qin, J., Shi, A., Nice, L., Bao, Y., Lorenz, A., et al. (2019). Genome-wide association study and genomic selection for soybean chlorophyll content associated with soybean cyst nematode. *BMC Genomics* 20:904. doi: 10.1186/s12864-019-6275-z
- Ravelombola, W., Qin, J., Shi, A., Nice, L., Bao, Y., Lorenz, A., et al. (2020). Genome-wide association study and genomic selection for tolerance of soybean biomass reduction under to soybean cyst nematode infestation. *PLoS ONE* 15:e0235089. doi: 10.1371/journal.pone.0235089
- Ravelombola, W., Shi, A., and Huynh, B. (2021). Loci discovery, network-guided approach, and genomic prediction for drought tolerance index in a multi-parent advanced generation intercross (MAGIC) cowpea population. *Hortic. Res.* 8:24. doi: 10.1038/s41438-021-00462-w
- Rendón-Anaya, M., Montero-Vargas, J. M., Saburido-Alvarez, S., Vlasova, A., Capella-Gutiérrez, S., Ordaz-Ortiz, J. J., et al. (2017). Genomic history of the origin and domestications of common bean in the Americas unveils its closest sister species. *Genome Biol.* 18:60. doi: 10.1186/s13059-017-1190-6
- Riggs, R. D., and Schmitt, D. P. (1988). Complete characterization of the race scheme for *Heterodera glycines*. *J. Nematol.* 20, 392–395.
- Schmutz, J., McClean, P., Mamidi, S., Wu, G., Cannon, S., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008
- Shi, A., Buckley, B., Mou, B. Q., Motes, D., Morris, J. B., Ma, J. B., et al. (2016). Association analysis of cowpea bacterial blight resistance in USDA cowpea germplasm. *Euphytica* 208, 143–155. doi: 10.1007/s10681-015-1610-1
- Shi, A., Qin, J., Mou, B. Q., Correll, J., Weng, Y. J., Brenner, D., et al. (2017). Genetic diversity and population structure analysis of spinach by single-nucleotide polymorphisms identified through genotyping-by-sequencing. *PLoS ONE* 12:e0188745. doi: 10.1371/journal.pone.0188745
- 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
- Smith, J. R., and Young, L. D. (2003). Host suitability of diverse lines of *Phaseolus vulgaris* to multiple populations of *Heterodera glycines*. *J. Nematol.* 35, 23–28.
- Song, Q., Jia, G., Hyten, D. L., Jenkins, J., Hwang, E.-Y., Schroeder, S. G., et al. (2015). SNP assay development for linkage map construction, anchoring whole-genome sequence, and other genetic and genomic applications in common bean. *G3 Genes Genomics Genet.* 5, 2285–2290. doi: 10.1534/g3.115.020594
- Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redoña, E., Jannink, J. L., et al. (2016). Genome-wide prediction models that incorporate *de novo* GWAS are a powerful new tool for tropical rice improvement. *Heredity* 116:395–408. doi: 10.1038/hdy.2015.113
- Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B., and Ingvarsson, P. K. (2017). Evaluating the accuracy of genomic prediction of growth and wood traits in two *Eucalyptus* species and their F1 hybrids. *BMC Plant Biol.* 17:110. doi: 10.1186/s12870-017-1059-6
- Tylka, G., and Marett, C. C. (2017). Known distribution of the soybean cyst nematode, *Heterodera glycines*, in the United States and Canada, 1954–2017. *Plant Health Prog.* 18:168. doi: 10.1094/PHP-05-17-0031-BR
- Valentini, G., Goncalves-Vidigal, M. C., Hurtado-Gonzales, O. P., Castro, S. A. D., Cregan, P. B., Song, Q. J., et al. (2017). High-resolution mapping reveals linkage between genes in common bean cultivar Ouro Negro conferring resistance to the rust, anthracnose, and angular leaf spot diseases. *Theor. Appl. Genet.* 130, 1705–1722. doi: 10.1007/s00122-017-2920-6
- Vlasova, A., Capella-Gutiérrez, S., Rendón-Anaya, M., Hernández-Onate, M., Minoche, A. E., et al. (2016). Genome and transcriptome analysis of

- the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes. *Genome Biol.* 17:32. doi: 10.1186/s13059-016-0883-6
- Wang, J., and Zhang, Z. (2020). GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *bioRxiv [Preprint]*. doi: 10.1101/2020.11.29.403170
- Wang, Q., Tian, F., Pan, Y., Buckler, E. S., and Zhang, Z. (2014). A SUPER powerful method for genome wide association study. *PLoS ONE* 9:e107684. doi: 10.1371/journal.pone.0107684
- Wen, L. W., Chang, H. X., Brown, P. J., Domier, L. L., and Hartman, G. L. (2019). Genome-wide association and genomic prediction identifies soybean cyst nematode resistance in common bean including a syntenic region to soybean Rhg1 locus. *Hortic Res.* 6:9. doi: 10.1038/s41438-018-0085-3
- Xavier, A., Muir, W. M., and Rainey, K. M. (2016). Assessing predictive properties of genome-wide selection in soybeans. *G3 Genes Genomics Genet.* 6, 2611–2616. doi: 10.1534/g3.116.032268
- Xavier, A., and Rainey, K. M. (2020). Quantitative genomic dissection of soybean yield components. *G3 Genes Genomics Genet.* 10, 665–675. doi: 10.1534/g3.119.400896
- Xu, Y. B., and Crouch, J. H. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci.* 48, 391–407. doi: 10.2135/cropsci2007.04.0191
- Yan, G. P., Plaisance, A., Chowdhury, I., Baidoo, R., Upadhaya, A., Pasche, J., et al. (2017). First report of the soybean cyst nematode *Heterodera glycines* infecting dry bean (*Phaseolus vulgaris* L.) in a commercial field in Minnesota. *Plant Dis.* 101:391. doi: 10.1094/PDIS-09-16-1257-PDN
- 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. doi: 10.1038/ng.3596
- Zhang, H. Y., Li, C. Y., Davis, E. L., Wang, J. S., Griffin, J. D., Kofsky, J., et al. (2016). Genome-wide association study of resistance to soybean cyst nematode (*Heterodera glycines*) HG Type 2.5.7 in wild soybean (*Glycine soja*). *Front. Plant Sci.* 7:1214. doi: 10.3389/fpls.2016.01214
- Zhang, J. P., Song, Q. J., Cregan, P. B., and Jiang, G. L. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). *Theor. Appl. Genet.* 129, 117–130. doi: 10.1007/s00122-015-2614-x
- Zhang, X. C., Perez-Rodriguez, P., Burgueno, J., Olsen, M., Buckler, E., Atlin, G., et al. (2017). Rapid cycling genomic selection in a multiparental tropical maize population. *G3-Genes Genomics Genet.* 7, 2315–2326. doi: 10.1534/g3.117.043141
- Zhang, Z. W., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat Genet.* 42, 355–360. doi: 10.1038/ng.546
- Zhao, H., Li, Y., Petkowski, J., Kant, S., Hayden, M. J., and Daetwyler, H. D. (2021). Genomic prediction and genomic heritability of grain yield and its related traits in a safflower genebank collection. *Plant Genome* 14:e20064. doi: 10.1002/tpg2.20064

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 reviewer DZ declared a shared affiliation, with no collaboration, with one of the authors QS to the handling editor at the time of the review.

Copyright © 2021 Shi, Gepts, Song, Xiong, Michaels and Chen. 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 for Yield and Virus Disease Resistance of Cassava Varieties Developed Over the Last Eight Decades in Uganda

Francis Manze^{1,2*}, Patrick Rubaihayo¹, Alfred Ozimati², Paul Gibson¹, Williams Esuma², Anton Bua², Titus Alicai², Chris Omongo² and Robert S. Kawuki²

¹ Department of Agricultural Production, Makerere University, Kampala, Uganda, ² National Crops Resources Research Institute (NaCRRI), Kampala, Uganda

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Li Li,
University of New England, Australia
Robooni Tumuhimbise,
National Agricultural Research
Organisation, Uganda
Patrick Chiza Chikoti,
Zambia Agriculture Research Institute
(ZARI), Zambia

*Correspondence:

Francis Manze
manze francis92@gmail.com

Specialty section:

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

Received: 11 January 2021

Accepted: 28 May 2021

Published: 21 June 2021

Citation:

Manze F, Rubaihayo P, Ozimati A,
Gibson P, Esuma W, Bua A, Alicai T,
Omongo C and Kawuki RS (2021)
Genetic Gains for Yield and Virus
Disease Resistance of Cassava
Varieties Developed Over the Last
Eight Decades in Uganda.
Front. Plant Sci. 12:651992.
doi: 10.3389/fpls.2021.651992

Achieving food security for an ever-increasing human population requires faster development of improved varieties. To this end, assessment of genetic gain for key traits is important to inform breeding processes. Despite the improvements made to increase production and productivity of cassava in Uganda at research level, there has been limited effort to quantify associated genetic gains. Accordingly, a study was conducted in Uganda to assess whether or not genetic improvement was evident in selected cassava traits using cassava varieties that were released from 1940 to 2019. Thirty-two varieties developed during this period, were evaluated simultaneously in three major cassava production zones; central (Namulonge), eastern (Serere), and northern (Loro). Best linear unbiased predictors (BLUPs) of the genotypic value for each clone were obtained across environments and regressed on order of release year to estimate annual genetic gains. We observed that genetic trends were mostly quadratic. On average, cassava mosaic disease (CMD) resistance increased by 1.9% per year, while annual genetic improvements in harvest index (0.0%) and fresh root yield (−5 kg per ha or −0.03% per ha) were non-substantial. For cassava brown streak disease (CBSD) resistance breeding which was only initiated in 2003, average annual genetic gains for CBSD foliar and CBSD root necrosis resistances were 2.3% and 1.5%, respectively. It's evident that cassava breeding has largely focused on protecting yield against diseases. This underpins the need for simultaneous improvement of cassava for disease resistance and high yield for the crop to meet its current and futuristic demands for food and industry.

Keywords: cassava breeding, cassava brown streak disease, cassava mosaic disease, yield related traits, genetic progress

INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is a major staple crop in the tropics (Food and Agriculture Organization of the United Nations (FAOSTAT), 2019) owing to its transformative potential to spur economic growth, rural development and food security (Otekunrin and Sawicka, 2019). Indeed, over 60% of the world's cassava is produced in Africa (Food and Agriculture Organization of the United Nations (FAOSTAT), 2019), where its roots are processed into various forms (Shittu et al., 2016) to feed millions of people on a daily basis (Prakash, 2018). Within sub-Saharan Africa (SSA),

cassava is recognized as a choice crop for climate change adaptation, as it performs reasonably well under prolonged droughts and marginal soils (Orek et al., 2020). It is for these reasons that cassava features predominantly in strategic plans for agricultural development of most SSA countries.

It suffices to note that cassava breeding efforts in Africa only began around 1930s (Storey and Nichols, 1938). During then, cassava mosaic disease (CMD) was a major breeding objective, as it had attained epidemic status on the continent (Legg and Thresh, 2000). Accordingly, pioneer cassava breeding efforts were initiated at Amani Research Station, Tanzania, to combat CMD. That breeding work involved interspecific hybridizations which led to the development and dissemination of cassava clones that were resistant to both CMD and cassava bacterial blight (Ortiz and Nassar, 2007).

The successful development of CMD resistant clones at Amani spurred an Africa-wide cassava research program that was instated at the International Institute of Tropical Agriculture (IITA) in Nigeria by 1971 (Hahn et al., 1980). Selected germplasm from Latin America, Asia and East Africa, along with cultivars from West Africa were collected to commence systematic genetic improvement of cassava at IITA (Hahn et al., 1980). Through that work, several elite genotypes with multiple resistances to prevalent pests and diseases and good culinary qualities were developed and disseminated to national breeding programs in Africa (Manyong et al., 2000).

In Uganda, CMD resistant varieties sourced from Tanzania formed a major part of the cassava production system between 1940s and 1980s (Otim-Nape et al., 2001), with clones such as Magana, Nyaraboke, Alado-Alado, Njire-Red, and Bamunanika predominating production in that period (Otim-Nape et al., 2001). It should be noted that systematic cassava improvement in Uganda only started in the 1980s when a second wave of CMD caused by coinfection of African Cassava Mosaic Virus (ACMV) and the East African Cassava Mosaic Virus Uganda (EACMV-UG) emerged (Gibson et al., 1996; Patil and Fauquet, 2009). Subsequently, elite cassava clones combining yield and resistance to CMD were sourced from IITA and evaluated in Uganda to select those with durable CMD resistance. Through this process, some outstanding varieties including NASE 1, NASE 2, and NASE 3 were identified and promoted for production in the early 1990s (Ssemakula et al., 2000).

Released varieties were meant to be used for two main food products: “boiled or fried roots” that predominates central and western Uganda, and “flour-based meal” that predominates the eastern and northern parts of the country. As such, emphasis was initially placed on development of varieties characterized by high fresh root yield and dry matter content, multiple resistance to pests and diseases, starch quality, and low hydrogen cyanide (Ssemakula et al., 2000).

However, with the outbreak of cassava brown streak disease (CBSD) in early 2000s (Alicai et al., 2007), considerable efforts were diverted toward breeding for CBSD resistance, as the disease had then attained epidemic status and caused immense yield losses (Kawuki et al., 2016). CBSD damages the starch bearing part of cassava rendering it unfit for consumption, thereby causing huge economic losses and food insecurity (Hillocks et al.,

2001). In fact, from the time when CBSD attained epidemic status in Uganda, cassava production in the country declined drastically from 4.9 million tons (MT) in the 2000s to the current 2.6 MT (Food and Agriculture Organization of the United Nations (FAOSTAT), 2019).

Another notable change in the 2010s, was the consideration of gender and integration of preferred end-user quality traits in cassava breeding operations (Esuma et al., 2019; Iragaba et al., 2019). Currently, cassava breeding in Uganda is designed to enhance key traits that contribute toward increased resilience, nutrition and productivity for the benefit of stakeholders involved in the production-processing-marketing-consumption continuum.

Through these breeding efforts, 21 cassava varieties have been released between 1993 and 2015, and several other elite clones developed using genomic selection (Ozimati et al., 2019). Despite the improvements made to increase production and productivity of cassava in Uganda at research level, there has been limited effort to quantify associated genetic gains. Quantifying such gains would guide cassava breeding processes, especially now when the rapidly increasing population demands faster development and deployment of improved varieties. Therefore, the objective of this study was to determine the rate of genetic gain per year for cassava traits that have been selected for between 1940 and 2019 in Uganda.

MATERIALS AND METHODS

Plant Material

A total of 32 cassava varieties were used for this study (Table 1) and these were divided into four categories. Category one comprised local varieties; these arose from selections from Amani Research Program in Tanzania and were deployed for cultivation in Uganda between 1940s and 1980s. Category two comprised varieties introduced from IITA and released in Uganda in the 1990s to combat CMD. Category three comprised a combination of varieties from IITA and Uganda; these were majorly developed for CMD resistance in the 2000s. Lastly, category four comprised varieties and elite clones developed in the 2010s to combat CBSD epidemic. All varieties were sourced from the Root Crops Research Program at the National Crops Resources Research Institute (NaCRRI) in Uganda, and had been maintained in Ngetta (northern Uganda), which is known to have low pressure of CBSD (Pariyo et al., 2015; Alicai et al., 2019). Sourcing planting materials from low disease pressure sites was important to ensure high vigor and uniform establishment.

Description of Trial Environments

All varieties were evaluated simultaneously at three environments representing major cassava agro-ecologies in Uganda, and this was done during the period April 2019 to May 2020. These environments were: Namulonge (0.5232°N, 32.6158°E), Serere (033°26'48.0"E, 01°32'22.6"N), and Loro (32°28'E, 2°12'N). Namulonge is located in the Lake Victoria crescent at an altitude of 1163 m above sea level (asl), and is characterized by reddish sandy-clay loam soils (Fungo et al., 2011). Serere is located in

TABLE 1 | Summary of attributes and origin of varieties used for genetic gain assessment.

| Code | Variety | Remarks | Status | Year | Special attributes at development and release |
|------|-------------|--|-----------|------|---|
| 1 | NASE 1 | Introduced from IITA as TMS 60142 | Released | 1993 | CMDt, high DMC, and low HCN |
| 2 | NASE 2 | Introduced from IITA as TMS 30337 | Released | 1993 | CMDt, good LR, and low HCN |
| 3 | NASE 3 | Introduced from IITA as TMS 30572 | Released | 1993 | CMDt, CBSDt, good LR, and low HCN |
| 4 | NASE 4 | Introduction from IITA | Released | 1999 | CMDr and low HCN |
| 5 | NASE 5 | Introduction from IITA | Released | 1999 | CMDt and low HCN |
| 6 | NASE 6 | Introduced from IITA as TMS 4 (2) 1425 | Released | 1999 | CMDr and low HCN |
| 7 | NASE 9 | Introduced from IITA as 30555-17 | Released | 2003 | CMDt, CBSDs, and low HCN |
| 8 | NASE 11 | Introduced from IITA as 92/NA-2 | Released | 2003 | CMDt, CBSDs, good LR, LUS, and low HCN |
| 9 | NASE 12 | MH95/0414 | Released | 2003 | CMDr, CBSDs, low HCN, and desirable CQ |
| 10 | NASE 13 | MH97/2961 | Released | 2011 | CMDr, CBSDs, high DMC, low HCN, and desirable CQ |
| 11 | NASE 14 | MM96/4271 | Released | 2011 | CMDr, CBSDt, high DMC, low HCN, and desirable CQ |
| 12 | NASE 15 | Derivative of TME14 | Released | 2011 | CMDr, CBSDt, high DMC, low HCN, and desirable CQ |
| 13 | NASE 16 | Derivative of Bamunanika | Released | 2011 | CMDr, CBSDs, high DMC, low HCN, and desirable CQ |
| 14 | NASE 18 | Derivative of TME14 | Released | 2011 | CMDr, CBSDt, high DMC, low HCN, and desirable CQ |
| 15 | NASE 19 | Derivative of TME14 | Released | 2011 | CMDr, CBSDt, high DMC, low HCN, and desirable CQ |
| 16 | NAROCASS 1 | NDL90/34HS | Released | 2015 | CMDr, CBSDt, high DMC, low HCN, and desirable CQ |
| 17 | NAROCASS 2 | Introduced from Tanzania as MM06130 | Released | 2015 | CMDr, CBSDt, high DMC, and desirable CQ |
| 18 | UG120124 | MM96/4271//MH04/2767 | Candidate | 2019 | CMDr, CBSDt, high DMC, and low HCN |
| 19 | UG110166 | Introduction from Tanzania | Candidate | 2019 | CMDr, CBSDt, high DMC, and low HCN |
| 20 | UG120024 | NASE 14/UG110043 | Candidate | 2019 | CMDr, CBSDt, high DMC, and low HCN |
| 21 | UG120156 | Introduction from Tanzania | Candidate | 2019 | CMDr, CBSDt, high DMC, low HCN, and high RWF |
| 22 | UG120183 | Introduction from Tanzania | Candidate | 2019 | CMDr, CBSDt, high DMC, and low HCN |
| 23 | UG120198 | Introduction from Tanzania | Candidate | 2019 | CMDr, CBSDt, high DMC, and low HCN |
| 24 | UG120193 | Introduction from Tanzania | Candidate | 2019 | CMDr, CBSDt, high DMC, low HCN, and high RWF |
| 25 | UG110164 | Introduction from Tanzania | Candidate | 2019 | CMDr, CBSDt, high DMC, and low HCN |
| 26 | Magana | Introduction from Tanzania | Landrace | 1940 | CMDt, quality flour and brew (popular in eastern Uganda) |
| 27 | Njire Red | Introduction from Tanzania | Landrace | 1940 | CMDt, soft when boiled or fried (popular in central Uganda) |
| 28 | Alado Alado | Introduction from Tanzania | Landrace | 1940 | CMDt, quality flour and brew (popular in northern Uganda) |
| 29 | Bamunanika | Introduction from Tanzania | Landrace | 1940 | CMDt, soft when boiled or fried (popular in central Uganda) |
| 30 | Bao | Introduction from Tanzania | Landrace | 1940 | CMDt, quality flour and brew (popular in northern Uganda) |
| 31 | Omo | Introduction from Tanzania | Landrace | 1940 | CMDt, EM, sweet, quality flour and brew (popular in west Nile) |
| 32 | Nyaraboke | Introduction from Tanzania | Landrace | 1940 | CMDt, soft when boiled or fried (popular in mid-western Uganda) |

IITA, International Institute of Tropical Agriculture; *CMDt*, tolerant to cassava mosaic disease; *CMDr*, resistant to cassava mosaic disease; *CBSDt*, tolerant to cassava brown streak disease; *CBSDs*, susceptible to cassava brown streak disease; *DMC*, dry matter content; *RWF*, resistant to whitefly; *LR*, leaf retention; *LUS*, long underground storage; *CQ*, culinary qualities; *HCN*, hydrogen cyanide; *EM*, early maturity. The candidate varieties (2019) have high CMD resistance, high DMC and high tolerance to CBSD.

the semi-arid zones of eastern Uganda at an altitude of 1085 m asl with sandy loamy soils (Isabirye et al., 2004). Loro, on the other hand, has an altitude of 1063 m asl is also characterized by sandy loamy soils (Isabirye et al., 2004). Namulonge and Serere were specifically chosen because they are known to have high disease pressure for CMD and CBSD as well as high whitefly (vector) populations (Alicai et al., 2019). Loro was considered a suitable site for yield assessment owing to low disease pressure and vector populations for CBSD (Pariyo et al., 2015). The rainfall distribution at the three trial sites is bimodal with peaks in March to May and August to October, and mean annual precipitation ranges between 500 and 2800 mm, while temperature ranges between 15°C and 30°C (Nsubuga et al., 2014).

Trial Design and Management

Trials at each site were planted in a randomized complete block design with two replications. Each clone was planted in five rows

of six plants at 1 x 1 m spacing, making a plot size of 20 m² with 30 plants. Adjacent plots were separated by 2-meter alleys to limit vegetative competition between varieties. Planting was done during the first growing season of 2019 (April) to ensure adequate soil moisture for sprouting. At 2 months after planting (MAP), six plants per plot were side-grafted with scions from highly infected TME 204, a standard CBSD susceptible check, to augment disease pressure for CBSD (Wagaba et al., 2013) at the three environments. All trials were conducted under standard agronomic practices for cassava (IITA, 1990).

Data Collection

Data on disease incidence and severity for CMD and CBSD were collected on each plant in a plot at 3 and 6 MAP. CMD severity was assessed on a scale of 1–5; where 1 = no visible disease symptoms, 2 = mild chlorotic pattern on entire leaflets or mild distortion at base of leaflets, rest of leaflets appearing green and

healthy, 3 = strong mosaic pattern on entire leaf, and narrowing and distortion of lower one-third of leaflets, 4 = severe mosaic, distortion of two-thirds of leaflets and general reduction of leaf size, and 5 = severe mosaic, distortion of four-fifths or more of leaflets, twisted and misshapen leaves (IITA, 1990). Similarly, CBSD foliar severity (CBSDfs) was scored on scale of 1–5, where 1 = no apparent symptoms, 2 = slight foliar chlorosis but with no stem lesions, 3 = pronounced foliar chlorosis and mild stem lesions with no die back, 4 = severe foliar chlorosis and severe stem lesions with no die back, and 5 = defoliation, severe stem lesions and die back (Gondwe et al., 2003).

At 12 MAP, trials were harvested to enable assessment of yield and other root attributes. All twelve plants within the net plot were harvested and partitioned into roots and above-ground biomass. Fresh root weight (FRW) and above-ground biomass were separately measured (kg plot^{-1}) using a hanging weighing scale of 200 kg capacity. Harvest index (HI) was calculated as the ratio of FRW to total plant biomass as described by Kawano et al. (1978). Fresh root yield (FRY) (tones ha^{-1}) was estimated by extrapolation of net plot root yields (Tumuhimbise et al., 2014). Root dry matter content (DMC) was determined by oven-drying of 100 g fresh samples at 80°C for 48 h, as described by Kawano et al. (1987). Lastly, data on cassava brown streak disease root necrosis incidence (CBSDri), and severity (CBSDrs) was recorded on all harvested roots/plot. Data on CBSDrs was collected using a standard scale of 1–5; where 1 = no observable necrosis, 2 = $\leq 5\%$ of root necrotic, 3 = 6 to 25% of root necrotic, 4 = 26 to 50% of root necrotic with mild root constriction, and while 5 showed greater than 50% of root necrosis with severe root constriction (Gondwe et al., 2003).

Data Analysis

For all measured traits, associated variance components were estimated by restricted maximum likelihood (Spilke et al., 2005). Effects of replicate nested in environment, variety, environment and variety by environment interaction were considered random, following the model below that was fitted using the *lmer* function in *lme4* package (Bates et al., 2015) in R (R Core Team, 2019).

$$Y_{ijk} = \mu + (R_j)E_k + V_i + E_k + VxE_{ik} + e_{ijk}$$

where, Y_{ijk} = phenotypic value; μ overall mean; $(R_j)E_k$ = random effect of replicate j nested in k^{th} environment such that $R_j \sim N(0, \sigma^2_j)$; V_i = random effect of the i^{th} variety with $V_i \sim N(0, \sigma^2_i)$; E_k = random effect of k^{th} environment with $E_k \sim N(0, \sigma^2_k)$; VxE_{ik} = random interaction effect of i^{th} variety with k^{th} environment such that $VxE_{ik} \sim N(0, \sigma^2_{ik})$; and e_{ijk} random residual that is assumed to be normally distributed with mean zero and variance σ^2 . Respective broad sense heritability (H^2) for each trait across environments was computed as:

$$H^2 = \frac{\sigma^2_V}{\sigma^2_V + \frac{\sigma^2_{VxE}}{n} + \frac{\sigma^2_e}{rn}}$$

Where, σ^2_V the variance component for variety; σ^2_{VxE} = the variance for variety by environment interaction; σ^2_e = the error variance; n = the number of environments; and r = the number

of replications. Accordingly, best linear unbiased predictors (BLUPs) for each variety were extracted using the *ranef* function in *lme4* package (Bates et al., 2015). Eventually, BLUPs were used to perform correlation analyses, compute selection index and estimate annual genetic gains for evaluated traits, as they provide better estimates of genotype performance for unbalanced datasets than fixed clone effects (Piepho et al., 2008).

A weight-free rank summation index (RSI) (Hallauer et al., 1988; Badu-Apraku et al., 2013) was used to rank variety performances based on nine traits: FRY, HI, DMC, CMDs, cassava mosaic disease incidence (CMDi), CBSDfi, CBSDfs, CBSDri, and CBSDrS. To estimate genetic gains, BLUPs were assigned to the year when the variety was released i.e., varieties specifically released in 1940, 1993, 1999, 2003, 2011, 2015, and the current candidate varieties of 2019. Because released years were unevenly distributed, traits were regressed on order of release year i.e., 1, 2, 3, 4, 5, 6, and 7 representing 1940, 1993, 1999, 2003, 2011, 2015, and 2019, respectively. Effects of order of release year were tested for linear and quadratic responses of evaluated traits by orthogonal polynomial contrasts to determine the model that would best fit the set of data for a specific trait.

Absolute gain for linear relationships was obtained following the statistical model: $y = a + bx$, where; y is dependent variable; x = independent variable (order of released year); a = intercept; and b regression slope, which is the absolute genetic gain per released order (de Felipe et al., 2016). The slope was thereafter divided by the number of years for the respective breeding period to determine the annual genetic gain. Relative gain was obtained by dividing the absolute annual gain by mean trait performance of oldest released year that served as the check.

For quadratic relationships, absolute annual gain was calculated as the slope between two released orders i.e., between 1 (1940) and 2 (1993), 2 (1993) and 3 (1999), and 3 (1999) and 4 (2003), etc divided by the number of years for the respective breeding period. Relative gain was obtained by dividing the absolute annual gain by mean trait performance of older released year for each specific breeding period. Thus genetic gains were assessed sequentially in phases and as an average.

RESULTS

Trait Heritabilities

Diseases (CMD and CBSD) and yield traits (DMC, HI, and FRYD) were differently affected by environment and genotypic effects (Table 2). For example, variety effects explained up to 96.4% of the total variance for CMD severity, while $< 20\%$ of total variance could be attributed to varieties for HI, DMC, and FRY. Indeed, highest heritability was registered for CMD ($H^2 = 0.96$) and lowest registered for harvest index ($H^2 = 0.43$). Overall, modest-high heritabilities i.e., $H^2 > 0.4$ were observed for all evaluated traits (Table 2).

Performance of Varieties Based on Rank Summation Index

Based on RSI, the top performers were mostly candidate clones (UG120024, UG120193, UG120183, UG120198, and UG110164),

TABLE 2 | Percentage of the total variance attributed to variety, environment and variety by environment interaction for evaluated traits.

| Source of variation | CMDi | CMDs | CBSDfi | CBSDfs | CBSDri | CBSDrs | DMC | HI | FRY |
|------------------------------------|------|------|--------|--------|--------|--------|------|------|------|
| Replicate/Environment | 0.0 | 0.5 | 11.4 | 11.2 | 0.0 | 6.1 | 4.3 | 0.0 | 11.8 |
| Variety | 95.4 | 96.4 | 33.9 | 35.2 | 45.7 | 36.5 | 18.2 | 5.8 | 13.6 |
| Environment | 0.4 | 0.0 | 43.6 | 39.2 | 22.1 | 12.3 | 60.3 | 86.7 | 63.0 |
| Variety*Environment | 3.1 | 2.7 | 7.6 | 10.8 | 21.9 | 30.4 | 11.5 | 3.5 | 6.1 |
| Residual | 1.1 | 0.4 | 3.5 | 3.6 | 10.3 | 14.7 | 5.7 | 4.0 | 5.5 |
| Genotype/Genotype*Environment | 30.8 | 35.7 | 4.4 | 3.2 | 2.1 | 1.2 | 1.6 | 1.6 | 2.2 |
| Broad-sense heritability (H^2) | 0.95 | 0.96 | 0.75 | 0.70 | 0.59 | 0.44 | 0.51 | 0.43 | 0.53 |

CMDi, cassava mosaic disease incidence at 6 months after planting; CMDs, cassava mosaic disease severity at 6 months after planting; CBSDFi, cassava brown streak disease foliar incidence at 6 months after planting; CBSDFs, cassava brown streak disease foliar severity at 6 months after planting; CBSDRi, cassava brown streak disease root incidence at 12 months after planting; CBSDRs, cassava brown streak disease root severity at 12 months after planting; DMC, root dry matter content; HI, Harvest index; FRY, Fresh root yield. Analysis based on data collected in 2019 at three sites; Namulonge (central region), Serere (eastern region) and Loro (northern region).

and varieties officially released in 2011 (NASE 15), 1999 (NASE 4), 2015 (NAROCASS 1 and NAROCASS 2) and NASE 1 (1993) (Table 3). On the other hand, worst performers mostly comprised of popular local varieties (Magana, Nyaraboke, Bamunanika and Njire Red), and varieties released in 1993 (NASE 2 and NASE 3) or 2011 (NASE 13 and NASE 14). Both local varieties and varieties released in 1990s exhibited higher CMD and CBD susceptibility compared to 2019 candidate clones or varieties released in 2015 (Tables 3, 4). Although varieties released in 2011 were generally resistant to CMD (incidence of $\leq 2.2\%$), they were susceptible to CBD (severity ≥ 2 and incidence $\geq 41\%$). Candidate clones exhibited high tolerance/resistance to CMD and CBD as well as high DMC (Tables 3, 4).

Genetic Gains for Disease Resistance and Yield Related Traits

Cassava mosaic disease severity correlated negatively and significantly with order of release year ($r = -0.9$, $P < 0.001$) (Table 5). However, there was a negative non-significant correlation between CMD severity and CBD foliar severity ($r = -0.36$, $P = 0.13$), and between CMD severity and CBD root necrosis severity ($r = -0.40$, $P = 0.08$). CMD severity reduced from mean severity score of 3.5 (varieties released in 1940) to 1.3 (candidate clones of 2019), attaining an average annual genetic gain of 1.9% (Table 6). Highest annual gains were registered for 1993 to 1999 (5.2%) and 1999 to 2003 (8.1%). However, CMD susceptibility increased by 4.5% per year from 2015 to 2019.

For CBD, we observed negative significant correlations between order of release year and CBD foliar severity ($r = -0.74$, $P < 0.001$). Similar observations were made between order of release year and CBD root necrosis severity ($r = -0.63$, $P < 0.01$) (Table 5). CBD foliar severity correlated positively and significantly with CBD root necrosis severity ($r = 0.67$, $P < 0.01$). CBD foliar severity reduced from symptom severity score of 2.1 in 2003 to 1.3 in 2019 and thus attaining an average annual genetic gain of 2.3% (Table 6). Highest annual gains were recorded for 2015 to 2019 (4.1% per year). Similarly, CBD root necrosis severity reduced from root necrosis score of 2.1 in 2003 to root necrosis score of 1.4 in 2019 and thus attaining an average annual genetic gain of 1.5%.

Much as order of release year correlated positively and significantly with dry matter content ($r = 0.40$, $P = 0.02$), we observed small, positive, nonsignificant correlations between order of release year and harvest index ($r = -0.11$, $P = 0.53$), plus fresh root yield ($r = 0.07$, $P = 0.73$). Fresh root yield correlated positively and significantly with harvest index ($r = 0.58$, $P < 0.001$). From 1940 to 2019, root dry matter content increased linearly from 37.6 to 39.4% with a genetic gain of 0.1% per year. Fresh root yield increased from 17.1 tons per ha in 1940 to 25.5 tons per ha in 1999 with an average annual gain of 0.06%. However, fresh root yield reduced from 25.6 tons/ha in 2003 to 18.6 tons/ha in 2019 at a rate of 0.12% per year (Table 6). Meanwhile, there were no genetic gains for harvest index between 1940 and 2019 (Figure 1).

DISCUSSION

Development and deployment of nutritious, stress-resilient, and high yielding cassava varieties requires identification and introgression of desirable alleles. As part of this process, routine assessment of genetic gain for key traits is necessary to identify gaps and quantify progress made toward attainment of prior defined breeding targets. Among the various methods for genetic gain assessment, growing released varieties in a common set of environments and regressing their trait means on year of release has gained popularity, as germplasm from recurrent selection programs is rarely available in breeding programs (Rutkoski, 2019). Accordingly, in this study, cassava varieties developed in Uganda between 1940 and 2019, were evaluated in 2019 to get insights into annual genetic gains. This was the first attempt to estimate genetic gain for selected cassava traits in Uganda.

Significant genotype variances were observed for all evaluated traits and thus, positively confirming the appreciable genetic variability in the evaluated clones and varieties (Ssemakula et al., 2000). The high heritabilities observed for disease traits are comparable to heritability estimates by Okul et al. (2018) and suggest that Namulonge (central Uganda) and Serere (eastern Uganda) are areas of high disease pressure for CMD and CBD. Consistency in variety or clone rankings for CMD resistance could imply durability of resistance in the tested genotypes.

TABLE 3 | Overall performance of clones based on rank summation index.

| Genotype | Year of release | CMDi | CMDs | CBSDfs | CBSDfi | CBSDrs | CBSDri | DMC | HI | FRY | RSI | Rank |
|-------------|-----------------|------|------|--------|--------|--------|--------|-----|----|-----|-----|------|
| Alado Alado | 1940 | 29 | 30 | 10 | 12 | 4 | 5 | 31 | 12 | 21 | 154 | 18 |
| Bao | 1940 | 28 | 31 | 6 | 6 | 20 | 17 | 30 | 20 | 18 | 176 | 22 |
| Omo | 1940 | 32 | 29 | 9 | 8 | 23 | 27 | 9 | 23 | 12 | 172 | 21 |
| Bamunanika | 1940 | 25 | 27 | 21 | 17 | 30 | 29 | 21 | 6 | 11 | 187 | 27 |
| Njire Red | 1940 | 30 | 32 | 29 | 28 | 10 | 11 | 13 | 8 | 20 | 181 | 23 |
| Nyaraboke | 1940 | 31 | 28 | 30 | 31 | 9 | 7 | 23 | 28 | 32 | 219 | 32 |
| Magana | 1940 | 27 | 24 | 23 | 26 | 17 | 19 | 11 | 28 | 30 | 205 | 30 |
| NASE 1 | 1993 | 26 | 25 | 11 | 11 | 1 | 1 | 18 | 5 | 25 | 123 | 9 |
| NASE 2 | 1993 | 21 | 22 | 26 | 22 | 21 | 26 | 24 | 13 | 8 | 183 | 24 |
| NASE 3 | 1993 | 23 | 23 | 22 | 20 | 22 | 24 | 26 | 24 | 27 | 211 | 31 |
| NASE 4 | 1999 | 12 | 9 | 15 | 19 | 4 | 6 | 28 | 2 | 5 | 100 | 6 |
| NASE 5 | 1999 | 24 | 26 | 17 | 18 | 23 | 21 | 25 | 1 | 2 | 157 | 19 |
| NASE 6 | 1999 | 13 | 15 | 31 | 27 | 16 | 20 | 27 | 9 | 26 | 184 | 26 |
| NASE 9 | 2003 | 19 | 19 | 28 | 25 | 15 | 23 | 19 | 22 | 13 | 183 | 24 |
| NASE 11 | 2003 | 20 | 20 | 14 | 14 | 25 | 22 | 10 | 4 | 1 | 130 | 11 |
| NASE 12 | 2003 | 2 | 9 | 27 | 30 | 18 | 15 | 20 | 19 | 23 | 163 | 20 |
| NASE 13 | 2011 | 10 | 11 | 32 | 32 | 32 | 32 | 5 | 30 | 10 | 194 | 29 |
| NASE 14 | 2011 | 2 | 4 | 24 | 29 | 31 | 31 | 15 | 32 | 24 | 192 | 28 |
| NASE 15 | 2011 | 2 | 5 | 13 | 13 | 19 | 16 | 6 | 20 | 4 | 98 | 4 |
| NASE 16 | 2011 | 2 | 1 | 18 | 16 | 26 | 28 | 17 | 18 | 6 | 132 | 12 |
| NASE 18 | 2011 | 1 | 3 | 25 | 24 | 29 | 30 | 7 | 15 | 14 | 148 | 17 |
| NASE 19 | 2011 | 2 | 5 | 19 | 21 | 28 | 25 | 11 | 14 | 15 | 140 | 14 |
| NAROCASS 1 | 2015 | 15 | 13 | 16 | 15 | 11 | 10 | 16 | 3 | 3 | 102 | 7 |
| NAROCASS 2 | 2015 | 8 | 5 | 8 | 10 | 12 | 14 | 14 | 25 | 31 | 127 | 10 |
| UG120193 | 2019 | 18 | 18 | 3 | 3 | 4 | 8 | 2 | 16 | 7 | 79 | 2 |
| UG120024 | 2019 | 9 | 1 | 6 | 7 | 3 | 3 | 3 | 9 | 28 | 69 | 1 |
| UG110164 | 2019 | 16 | 16 | 5 | 5 | 13 | 13 | 22 | 6 | 9 | 105 | 8 |
| UG120183 | 2019 | 17 | 17 | 1 | 1 | 4 | 4 | 4 | 16 | 16 | 80 | 3 |
| UG120124 | 2019 | 2 | 5 | 4 | 4 | 26 | 18 | 29 | 31 | 22 | 141 | 15 |
| UG120156 | 2019 | 11 | 12 | 12 | 9 | 14 | 12 | 8 | 26 | 28 | 132 | 12 |
| UG120198 | 2019 | 22 | 21 | 2 | 2 | 2 | 2 | 1 | 27 | 19 | 98 | 4 |
| UG110166 | 2019 | 14 | 13 | 20 | 23 | 8 | 9 | 32 | 11 | 17 | 147 | 16 |

CMDi, cassava mosaic disease incidence at 6 months after planting; CMDs, cassava mosaic disease severity at 6 months after planting; CBSDfi, cassava brown streak disease foliar incidence at 6 months after planting; CBSDfs, cassava brown streak disease foliar severity at 6 months after planting; CBSDri, cassava brown streak disease root incidence at 12 months after planting; CBSDrs, cassava brown streak disease root severity at 12 months after planting; DMC, root dry matter content; HI, Harvest index; FRY, Fresh root yield; RSI, rank summation index. BLUPs for disease traits (CMD and CBSD) based only on data from Namulonge and Serere owing to low disease pressure at Loro. Genotypes were ranked based on their BLUP values for each trait.

TABLE 4 | Means for cassava traits selected for between 1940 and 2019 in Uganda.

| Year of release | No of Varieties | CMDs | DMC | HI | FRY | CBSDfs | CBSDrs |
|-----------------|-----------------|------------|------------|--------------|------------|-----------|-----------|
| 1940 | 7 | 3.5 ± 0.2 | 37.6 ± 1.1 | 0.37 ± 0.03 | 17.1 ± 3.5 | 1.8 ± 0.2 | 1.8 ± 0.3 |
| 1993 | 3 | 2.3 ± 0.3 | 37.4 ± 0.4 | 0.40 ± 0.003 | 20.9 ± 8.5 | 2.0 ± 0.3 | 1.8 ± 0.4 |
| 1999 | 3 | 1.6 ± 0.6 | 35.9 ± 0.6 | 0.45 ± 0.01 | 25.5 ± 7.4 | 2.1 ± 0.2 | 1.8 ± 0.3 |
| 2003 | 3 | 1.4 ± 0.2 | 38.5 ± 0.3 | 0.38 ± 0.01 | 25.6 ± 7.1 | 2.1 ± 0.2 | 2.1 ± 0.1 |
| 2011 | 6 | 1.1 ± 0.02 | 39.8 ± 0.6 | 0.30 ± 0.02 | 22.7 ± 2.5 | 2.2 ± 0.2 | 2.9 ± 0.2 |
| 2015 | 2 | 1.1 ± 0.1 | 38.8 ± 0.1 | 0.40 ± 0.07 | 17.5 ± 8.1 | 1.7 ± 0.5 | 1.5 ± 0.0 |
| 2019 | 8 | 1.3 ± 0.1 | 39.5 ± 1.3 | 0.30 ± 0.02 | 18.6 ± 1.8 | 1.3 ± 0.1 | 1.4 ± 0.2 |

CMDs, cassava mosaic disease severity; DMC, dry matter content; HI, harvest index; FRY, fresh root yield; CBSDfs, cassava brown streak disease foliar severity; CBSDrs, cassava brown streak disease root severity.

Indeed, local and varieties released in 1990s consistently registered higher CMD susceptibility when compared to recent elite clones or varieties released in 2015 (Table 3). However,

variety or clone rankings for CBSD resistance were not consistent across environments, possibly because there could be different cassava brown streak virus strains that are resident in test

TABLE 5 | Pearson's correlation coefficients among selected cassava traits evaluated in cassava varieties released between 1940 and 2019 in Uganda.

| | DMC | CMDs | CBSDfs | CBSDrs | HI | FRY |
|-----------------------|-------|---------|----------|---------|---------|------|
| CMD6s | -0.21 | 1 | | | | |
| CBSDfs | -0.29 | -0.36 | 1 | | | |
| CBSDrs | -0.17 | -0.40 | 0.71*** | 1 | | |
| HI | -0.21 | 0.07 | 0.03 | -0.34 | 1 | |
| FRY | 0.06 | -0.08 | 0.17 | 0.22 | 0.58*** | 1 |
| Order of release year | 0.40* | -0.9*** | -0.74*** | -0.63** | -0.11 | 0.07 |

CMD6s, cassava mosaic disease severity at 6 months after planting; CBSDfs, cassava brown streak disease severity at 6 months after planting; CBSDrs, cassava brown streak disease root necrosis severity at 12 months after planting; DMC, dry matter content; HI, harvest index; FRY, fresh root yield; * $P < 0.05$, ** $P < 0.01$; *** $P < 0.001$. Correlations including CBSDfs and CBSDrs were performed using varieties or clones developed between 2003 and 2019 because selection for CBSD resistance began in 2003. Two sets of correlations were performed: (1) between evaluated traits and order of release year, and (2) amongst the evaluated traits across the released years.

environments. Cassava brown streak viruses [cassava brown streak virus (CBSV) and Uganda cassava brown streak virus (UCBSV)] have been reported to evolve rapidly, a phenomenon that could influence virulence (Ndunguru et al., 2015; Alicai et al., 2016) and thus amplify genotype by environment interactions (Pariyo et al., 2015; Okul et al., 2018). These findings further underpin the need for systematic evaluation and screening for CBSD in locations that are truly hotspots so as to discern resistant from susceptible clones.

Generally, candidate varieties and recently released varieties exhibited higher disease resistance (Tables 3, 4). Indeed, some of the candidate varieties e.g., UG120156 and UG120024 have also been reported by Okul et al. (2018) to exhibit high CBSD resistance. One possible explanation for this is that these candidate clones and/or varieties were selected for dual resistances to CMD and CBSD, which was not the case with varieties released before 2011. An exceptional clone was NASE

4, a variety released in 1999, which ranked among the top 10 performers; its ability to maintain superior and stable performance over a wide range of environments could explain this trend (Adriko et al., 2011).

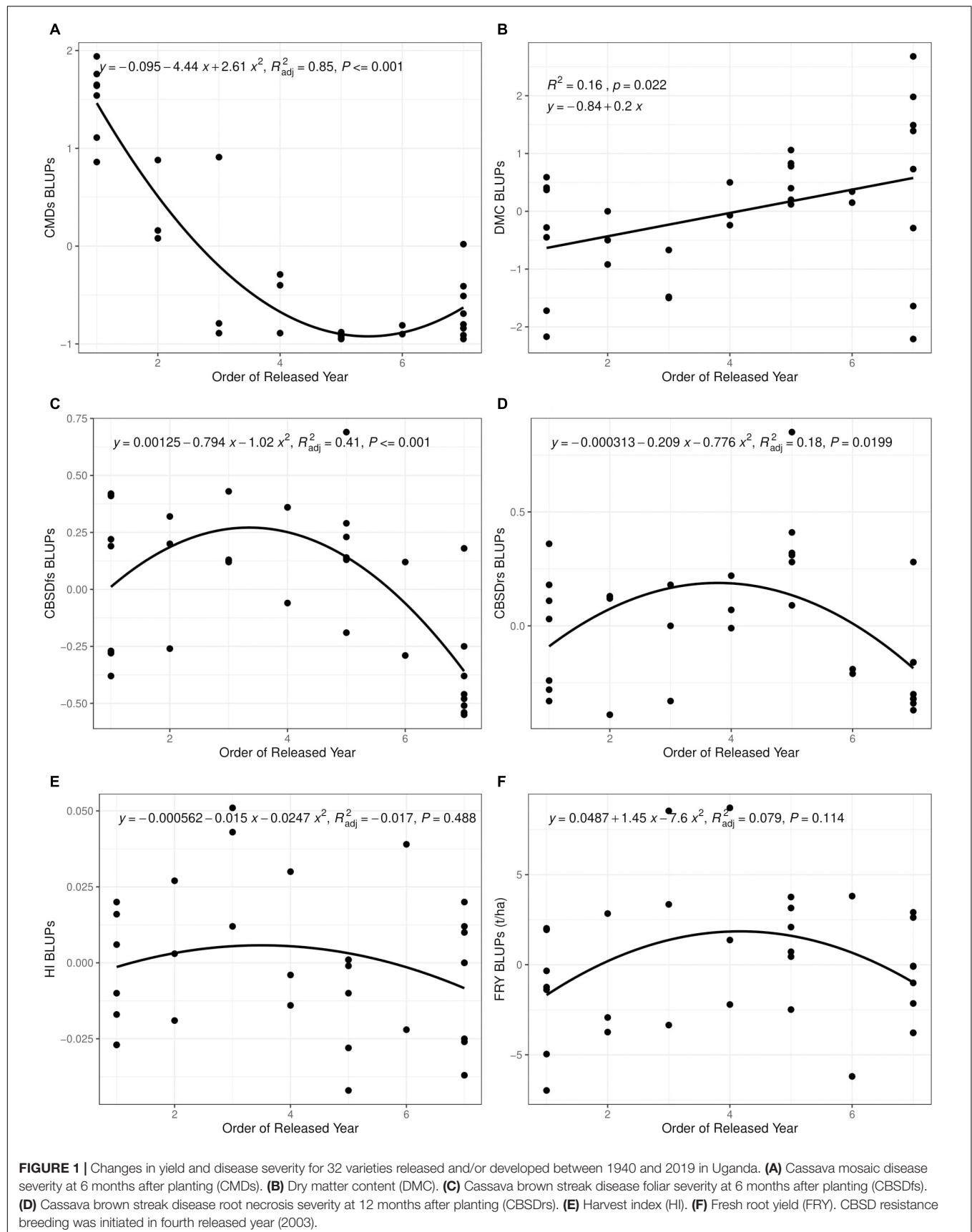
Local varieties such as Magana (popular in eastern region), Nyaraboke (popular in mid-western region), Bamunanika (popular in central region) and Njire Red (popular in central region), were among the worst performers. These varieties showed high susceptibility to both CMD and CBSD (Tables 3, 4). It is important to note that these local varieties were among the first CMD resistant clones developed in 1930s in Amani (Tanzania) and introduced into Uganda in the 1940s for cultivation (Legg and Thresh, 2000). These clones were deployed for production in 1950s and formed a major part of the cassava production system in Uganda until the 1980s (Otim-Nape et al., 2001), when a second wave of CMD caused by co-infection of African Cassava Mosaic Virus (ACMV) and the recombinant strain of the East African Cassava Mosaic Virus (EACMV-UG) emerged (Patil and Fauquet, 2009). The breakdown of CMD resistance in local varieties and varieties released in early 1990s (Table 3) is likely due to the long exposure to viruses or synergistic infections from the different cassava mosaic geminiviruses (CMGs).

Following the CBSD outbreak in Uganda in the early 2000s (Alicai et al., 2007), efforts were initiated to develop and release varieties that combine both CMD and CBSD resistance. The first batch of these varieties were officially released in 2011, all in an effort to limit spread and damage inflicted by CBSD. Notable of these were: NASE 14, NASE 15, NASE 16, NASE 18, and NASE 19. However, in the present study, these varieties maintained CMD resistance, but succumbed to CBSD, as exhibited in their respective CBSD foliar incidence (Table 4). Given that this assessment was done 8 years after these varieties were released, it is likely that the high root necrosis severity scores (Tables 3, 4) are a reflection of increased virus load accumulating in the vegetative tissues during this propagation

TABLE 6 | Genetic gains for cassava traits selected for between 1940 and 2019 in Uganda.

| Breeding period | | | Absolute annual gain | | | | | Relative annual gain (%) | | | | |
|--------------------------|------|--------------|----------------------|-------|--------|--------|--------|--------------------------|------|--------|--------|-------|
| From | To | No. of years | CMDs | DMC | CBSDfs | CBSDrs | FRY | CMD | DMC | CBSDfs | CBSDrs | FRY |
| 1940 | 1993 | 53 | -0.02 | 0.004 | 0.004 | 0.004 | 2.6 kg | -0.57 | 0.01 | 0.22 | 0.22 | 0.02 |
| 1993 | 1999 | 6 | -0.12 | 0.033 | 0.010 | 0.010 | 20 kg | -5.20 | 0.09 | 0.50 | 0.56 | 0.10 |
| 1999 | 2003 | 4 | -0.13 | 0.050 | 0.000 | 0.000 | 15 kg | -8.10 | 0.14 | 0.00 | 0.00 | 0.06 |
| 2003 | 2011 | 8 | -0.03 | 0.025 | -0.010 | -0.008 | 0.0 kg | -1.80 | 0.06 | -0.48 | -0.38 | 0.00 |
| 2011 | 2015 | 4 | 0.00 | 0.050 | -0.050 | -0.033 | -30 kg | 0.00 | 0.13 | -2.30 | -1.10 | -0.13 |
| 2015 | 2019 | 4 | 0.05 | 0.050 | -0.070 | -0.047 | -38 kg | 4.50 | 0.14 | -4.10 | -3.10 | -0.21 |
| Average genetic gain | | | -0.04 | 0.035 | -0.04 | -0.03 | -5 kg | -1.90 | 0.10 | -2.30 | -1.50 | -0.03 |
| Adjusted R^2 linear | | | 0.63 | 0.16 | 0.14 | 0.02 | 0.02 | | | | | |
| Adjusted R^2 quadratic | | | 0.85 | 0.10 | 0.41 | 0.18 | 0.079 | | | | | |

CMDs, cassava mosaic disease severity at 6 months; DMC, dry matter content; CBSDfs, cassava brown streak disease foliar severity at 6 months after planting; CBSDrs, cassava brown streak disease root severity at 12 months after planting; FRY, fresh root yield; R^2 , coefficient of determination of the relationship between order of release year and the changes in traits over the years. Average annual gains (absolute and relative) for resistance to CBSD were computed using estimates from 2003 to 2019 because selection for the trait only began in 2003. There were no genetic gains for harvest index between 1940 and 2019. With the exception of DMC where genetic gains were estimated using slope of linear regression, annual genetic gains for all other traits were estimated using the slope between two released orders from the quadratic graphs, because the quadratic model provided higher R^2 values.



period (Shirima et al., 2019). Similar observations were made by Mukiibi et al. (2018) and Okul et al. (2018), who reported that NASE 14 (released in 2011) registered high CBSD foliar and root incidence and severity after 6 years of release. This situation may be attributed to changes in the composition of virus species and/or virulence that overwhelms host defense systems and cause resistance breakdown or degeneration (Shirima et al., 2019). The clonal nature of cassava propagation amplifies this problem.

Correlation analyses were performed to assess relationships between order of release year and traits evaluated (Table 5). The significant linear relationships between order of release year and CMD resistance plus dry matter content suggest that breeding efforts between 1940 and 2019 were successful in developing CMD resistant genotypes with high dry matter content. Significant negative correlations between order of release year and CBSD resistance also suggest that breeding efforts undertaken between 2003 and 2019 majorly focused on development and/or release of CBSD resistant varieties. The small nonsignificant correlations between order of release year and yield-related traits (FRY and HI) are indicative of preferential selection and release of genotypes with more emphasis placed on disease resistances as compared yield.

Direct selection for disease resistance without similar efforts devoted to yield-traits could explain the non-significant positive correlations between CBSD resistance with FRY or HI. On the other hand, high significant positive correlation between CBSD foliar and CBSD root necrosis severity, could imply that both traits were directly selected for, as witnessed by their respective reductions across years of release. Negative nonsignificant correlations between CBSD resistance and CMD resistance between 2003 and 2019, could suggest that high levels of CMD resistance had been attained at the time when selection for CBSD resistance was initiated, and therefore, most of the clones were tolerant to CMD, but had not attained similar levels of resistance for CBSD.

Based on regression analyses, CMD severity reduced by an average of 1.9% per year between the period 1940 and 2019. This genetic gain estimate is higher than that provided by Okechukwu and Dixon (2008), who reported 0.65% genetic gain per year for CMD resistance among IITA clones developed in Nigeria between 1970 and 2000. The highly significant genetic gain per year for CMD resistance could be explained in three ways. Firstly, breeding efforts targeting CMD resistance have been ongoing since 1930s (Legg and Thresh, 2000), which is sufficient time for increasing the frequency of resistance alleles in the breeding population through recurrent selection (Hallauer et al., 1988). Secondly, CMD resistance is largely governed by additive genetic effects (Hahn et al., 1980; Wolfe et al., 2016; Rabbi et al., 2020), which makes it amenable to genetic gains from recurrent selection. Thirdly, that deployed CMD resistance was effective against the prevalent cassava mosaic geminiviruses. Indeed, latest findings by Mukiibi et al. (2018) have showed that both single and coinfection of ACMV and EACMV-UG do exist in Uganda. The 4.5% increase in CMD susceptibility between 2015 and 2019 may be attributed to tradeoffs during selection for combined resistance to CBSD and CMD or use of CBSD resistant parents that are deficient in CMD resistance.

Much as research efforts to combat CBSD began in early 2000s when the disease had attained epidemic status in Uganda (Alicai et al., 2007), some varieties like NASE 1 that were released in 1993, exhibited high CBSD tolerance (Table 3). This finding could indicate that CBSD resistance alleles were present in IITA germplasm, from which NASE 1 was derived. Since 2003 when systematic CBSD resistance improvement began, there were average genetic gains of 2.3% per year for CBSD foliar resistance, and 1.5% per year for CBSD root necrosis resistance (Table 6). These genetic gains for CBSD resistance within such a relatively short timeframe could be attributed to the concerted and systematic approaches taken to harness and utilize available genetic resources in cassava breeding (Abaca et al., 2012; Kaweesi et al., 2014; Pariyo et al., 2015; Kawuki et al., 2016; Okul et al., 2018; Ozimati et al., 2018). Predominance of additive gene effects for CBSD resistance (Kulembeka et al., 2012; Chipeta et al., 2018), which can be exploited through recurrent selection, have equally enabled consolidation of gains.

Between 1940 and 2019, generally 5 kg per ha per year were lost for fresh root yield and no genetic gains in harvest index were observed; equally low genetic gains were recorded for dry matter content (0.1% per year) (Table 6). This is contrary to findings from earlier studies by Okechukwu and Dixon (2008), and Ceballos et al. (2020), who reported annual genetic gains of 1.2% and 1.0% for fresh root yield in Nigeria and Thailand, respectively. Differences in selection strategies customized to address local needs in Uganda, Nigeria and Thailand could explain this variation. For example, breeding programs in South East Asia have for long, mainly focused on developing cassava clones with high yield and root quality traits such as starch (Ceballos et al., 2020). Similarly, cassava breeding programs in West Africa (Nigeria) have focused on development of cassava clones that combine high fresh root yield, root quality and CMD resistance (Manyong et al., 2000). In Uganda, however, critical traits selected for include; dual resistance to CMD and CBSD, high yield and desirable root quality (Kawuki et al., 2016). Certainly, selection for several traits limits genetic progress as it leads to compromising tradeoffs amongst target traits. For example, before CBSD emerged in Uganda, fresh root yield increased from 17.1 tons in 1940 to 25.6 tons in 2003 (Table 4). However, when CBSD attained epidemic status in the early 2000s, fresh root yield reduced from 25.6 tons in 2003 to 18.6 tons in 2019. Another good example is the sharp contrast between fresh root yield and CBSD resistance observed in clones UG120024 and UG120156 (Table 3).

CONCLUSION

The study described herein was conducted to estimate annual genetic gains for critical cassava traits that have been selected for between 1940 and 2019 in Uganda. Based on the generated datasets, this study revealed that there was significant annual genetic improvement of cassava for resistance to CMD and CBSD. Findings from the present study also demonstrated that the annual rate of genetic gain for cassava yield in Uganda is not sufficient to achieve the desired output necessary to reach the

cassava production demand predicted for 2050. This underpins the urgent need to incorporate simultaneous selection for disease resistance and high yield for the crop to meet its current and futuristic demands for food and industry.

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.

AUTHOR CONTRIBUTIONS

FM involved in data collection, analysis, and the manuscript writing. PR and PG involved in review and supervision. AO and WE involved in data collection, the manuscript editing, and review. AB and CO involved in review. TA involved in the manuscript review and editing. RK involved in data collection, analysis, the manuscript editing, and review and funding acquisition. All authors contributed to the article and approved the submitted version.

REFERENCES

- Abaca, A., Kawuki, R., Tukamuhabwa, P., Baguma, Y., Pariyo, A., Alicai, T., et al. (2012). Evaluation of local and elite *Cassava genotypes* for resistance to cassava brown streak disease in Uganda. *J. Agron.* 11, 65–72. doi: 10.3923/ja.2012.65.72
- Adriko, J., Sserubombwe, W. S., Adipala, E., Bua, A., Thresh, J. M., and Edema, R. (2011). Response of improved cassava varieties in Uganda to cassava mosaic disease (CMD) and their inherent resistance mechanisms. *Afr. J. Agric. Res.* 6, 521–531.
- Alicai, T., Ndunguru, J., Sseruwagi, P., Tairo, F., Okao-Okuja, G., Nanvubya, R., et al. (2016). Cassava brown streak virus has a rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. *Sci. Rep.* 6, 1–14. doi: 10.1038/srep36164
- Alicai, T., Omongo, C. A., Maruthi, M. N., Hillocks, R. J., Baguma, Y., Kawuki, R., et al. (2007). Re-emergence of cassava brown streak disease in Uganda. *Plant Dis.* 91, 24–29. doi: 10.1094/PD-91-0024
- Alicai, T., Szyniszewska, A. M., Omongo, C. A., Abidrabo, P., Okao-okuja, G., Baguma, Y., et al. (2019). Expansion of the cassava brown streak pandemic in Uganda revealed by annual field survey data for 2004 to 2017. *Sci. Data* 6, 1–8. doi: 10.1038/s41597-019-0334-9
- Badu-Apraku, B., Yallou, C. G., and Oyekunle, M. (2013). Genetic gains from selection for high grain yield and Striga resistance in early maturing maize cultivars of three breeding periods under Striga-infested and Striga-free environments. *Field Crops Res.* 147, 54–67. doi: 10.1016/j.fcr.2013.03.022
- Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Statist. Softw.* 67:51. doi: 10.18637/jss.v067.i01
- Ceballos, H., Rojanaridpiched, C., Phumichai, C., Becerra, L. A., Kittipadukul, P., Iglesias, C., et al. (2020). Excellence in cassava breeding: perspectives for the future. *Crop Breed Genet. Genom.* 2:e200008.
- Chipeta, M. M., Melis, R., and Shanahan, P. (2018). Gene action controlling cassava brown streak disease resistance and storage root yield in cassava. *Euphytica* 214, 1–15. doi: 10.1007/s10681-018-2196-1
- de Felipe, M., Gerde, J. A., and Rotundo, J. L. (2016). Soybean genetic gain in maturity groups III to V in Argentina from 1980 to 2015. *Crop Sci.* 56, 3066–3077. doi: 10.2135/cropsci2016.04.0214
- Esuma, W., Nanyonjo, A. R., Miir, R., Angudubo, S., and Kawuki, R. S. (2019). Men and women's perception of yellow-root cassava among rural farmers in eastern Uganda. *Agric. Food Security* 8, 1–9. doi: 10.1186/s40066-019-0253-1
- Food and Agriculture Organization of the United Nations (FAOSTAT) (2019). *FAOSTAT Statistical Database*. Rome: FAO.
- Fungo, B., Grunwald, S., Tenywa, M. M., and Vanlauwe, B. (2011). Lunnyu soils in the lake victoria basin of Uganda: link to toposequence and soil type. *Afr. J. Env. Sci. Technol.* 5, 15–24.
- Gibson, R. W., Legg, J. P., and Otim-Nape, G. W. (1996). Unusually severe symptoms are a characteristic of the current epidemic of mosaic virus disease of cassava in Uganda. *Ann. Appl. Biol.* 128, 479–490. doi: 10.1111/j.1744-7348.1996.tb07108.x
- Gondwe, F. M. T., Mahungu, N. M., Hillocks, R. J., Raya, M. D., Moyo, C. C., Soko, M. M., et al. (2003). "Economic losses experienced by small-scale farmers in Malawi due to cassava brown streak virus disease," in *Proceedings of the 1st International Workshop*, eds J. P. Legg and R. J. Hillocks (Kenya), 27–30.
- Hahn, S. K., Terry, E. R., and Leuschner, K. (1980). Breeding cassava for resistance to cassava mosaic disease. *Euphytica* 29, 673–683. doi: 10.1007/BF00023215
- Hallauer, A. R., Carena, M. J., and Filho, M. J. B. (1988). *Quantitative Genetics in Maize Breeding*. Berlin: Springer, doi: 10.1007/978-1-4419-0766-0
- Hillocks, R. J., Raya, M. D., Mtunda, K., and Kiozia, H. (2001). Effects of brown streak virus disease on yield and quality of cassava in Tanzania. *J. Phytopathol.* 149, 389–394. doi: 10.1046/j.1439-0434.2001.00641.x
- IITA. (1990). *Cassava in Tropical Africa. A Reference Manual*. United Kingdom: Chayce Publication Services.
- Iragaba, P., Nuwamanya, E., Wembabazi, E., Baguma, Y., Dufour, D., Earle, E. D., et al. (2019). Estimates for heritability and consumer-validation of a penetrometer method for phenotyping softness of cooked cassava roots. *Afr. Crop Sci. J.* 27, 147–163.
- Isabirye, M., Mwesige, D., Ssali, H., Magunda, M., and Lwasa, J. (2004). Soil resource information and linkages to agricultural production. *Uganda J. Agric. Sci.* 9, 215–221.
- Kawano, K., Daza, P., Amaya, A., Rios, M., and Goncalves, W. M. F. (1978). Evaluation of cassava germplasm for productivity. *Crop Sci.* 18, 377–380. doi: 10.2135/cropsci1978.0011183X001800030006x
- Kawano, K., Fukuda, W. M. G., and Cenpukdee, U. (1987). Genetic and environmental effects on dry matter content of cassava root. *Crop Sci.* 27, 69–74. doi: 10.2135/cropsci1987.0011183X002700010018x
- Kaweisi, T., Kawuki, R., Kyaligonza, V., Baguma, Y., Tusiime, G., and Ferguson, M. E. (2014). Field evaluation of selected cassava genotypes for cassava brown

FUNDING

Funding for this work was provided by Cornell University through a sub-award agreement (N0. 84941-11038) between NaCRRI and Cornell University. This work was funded by Cornell University through "Next Generation Cassava Breeding Project."

ACKNOWLEDGMENTS

Special thanks go to the field-based staff of NaCRRI's Root Crops Program for assistance in trial establishment, management, and data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.651992/full#supplementary-material>

- streak disease based on symptom expression and virus load. *Virology J.* 11:216. doi: 10.1186/s12985-014-0216-x
- Kawuki, R. S., Kweesi, T., Esuma, W., Pariyo, A., Kayondo, I. S., Ozimati, A., et al. (2016). Eleven years of breeding efforts to combat cassava brown streak disease. *Breed. Sci.* 66, 560–571. doi: 10.1270/jsbbs.16005
- Kulembeka, H. P., Ferguson, M., Herselman, L., Kanju, E., Mkamillo, G., Masumba, E., et al. (2012). Diallel analysis of field resistance to brown streak disease in cassava (*Manihot esculenta* Crantz) landraces from Tanzania. *Euphytica* 187, 277–288. doi: 10.1007/s10681-012-0730-0
- Legg, J. P., and Thresh, J. M. (2000). Cassava mosaic virus disease in East Africa: a dynamic disease in a changing environment. *Virus Res.* 71, 135–149. doi: 10.1016/S0168-1702(00)00194-5
- Manyong, V. M., Makinde, K. O., Bokanga, M., and Whyte, J. (2000). *The Contribution of IITA-Improved Cassava to Food Security in Sub-Saharan Africa: an Impact Study*. Nigeria: International Institute of Tropical Agriculture.
- Mukiibi, D. R., Alicai, T., Kawuki, R., Okao-Okuja, G., Tairo, F., Sseruwagi, P., et al. (2018). Resistance of advanced cassava breeding clones to infection by major viruses in Uganda resistance of advanced cassava breeding clones to infection by major viruses in Uganda. *Crop Protection* 115, 104–112. doi: 10.1016/j.cropro.2018.09.015
- Ndunguru, J., Sseruwagi, P., Tairo, F., Stomeo, F., Maina, S., Djinkeng, A., et al. (2015). Analyses of twelve new whole genome sequences of cassava brown streak viruses and ugandan cassava brown streak viruses from east africa: diversity, supercomputing and evidence for further speciation. *PLoS One* 10:e0141939. doi: 10.1371/journal.pone.0139321
- Nsubuga, F. W. N., Botai, O. J., Olwoch, J. M., Rautenbach, C. J., Bevis, Y., and Adetunji, A. O. (2014). The nature of rainfall in the main drainage sub-basins of Uganda. *Hydrol. Sci. J.* 59, 278–299. doi: 10.1080/02626667.2013.804188
- Okechukwu, R. U., and Dixon, A. G. O. (2008). Genetic gains from 30 years of cassava breeding in Nigeria for storage root yield and disease resistance in elite cassava genotypes. *J. Crop Improvement* 22, 181–208. doi: 10.1080/15427520802212506
- Okul, V. A., Ochwo-Ssemakula, M., Kweesi, T., Ozimati, A., Mrema, E., Mwale, E. S., et al. (2018). Plot based heritability estimates and categorization of cassava genotype response to cassava brown streak disease. *Crop Protection* 108, 39–46. doi: 10.1016/j.cropro.2018.02.008
- Orek, C., Gruissem, W., Ferguson, M., and Vanderschuren, H. (2020). Morphological and molecular evaluation of drought tolerance in cassava (*Manihot esculenta* Crantz). *Field Crops Res.* 255:107861. doi: 10.1016/j.fcr.2020.107861
- Ortiz, R., and Nassar, N. M. A. (eds) (2007). “Cassava improvement to enhance livelihoods in sub-saharan africa and northeastern brazil,” in *Proceeding of the First International Meeting on Cassava Breeding, Biotechnology and Ecology*.
- Otekunrin, O., and Sawicka, B. (2019). Cassava, a 21st century staple crop: how can nigeria harness its enormous trade potentials? *Acta Sci. Agric.* 3, 194–202. doi: 10.31080/asag.2019.03.0586
- Otim-Nape, G. W., Alicai, T., and Thresh, M. J. (2001). Changes in the incidence and severity of Cassava mosaic virus disease, varietal diversity and cassava production in Uganda. *Ann. Appl. Biol.* 138, 313–327.
- Ozimati, A., Kawuki, R., Esuma, W., Kayondo, S. I., Pariyo, A., Wolfe, M., et al. (2019). Genetic variation and trait correlations in an East African Cassava breeding population for genomic selection. *Crop Sci.* 59, 460–473. doi: 10.2135/cropsci2018.01.0060
- Ozimati, A., Kawuki, R., Esuma, W., Kayondo, I. S., Wolfe, M., Lozano, R., et al. (2018). Training population optimization for prediction of cassava brown streak disease resistance in west african clones. *G3-Genes Genom. Genet.* 12, 3903–3913. doi: 10.1534/g3.118.200710
- Pariyo, A., Baguma, Y., Alicai, T., Kawuki, R., Kanju, E., Bua, A., et al. (2015). Stability of resistance to cassava brown streak disease in major agro-ecologies of Uganda. *J. Plant Breed. Crop Sci.* 7, 67–78. doi: 10.5897/jpbcs2013.0490
- Patil, B. L., and Fauquet, C. M. (2009). Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol. Plant Pathol.* 10, 685–701. doi: 10.1111/j.1364-3703.2009.00559.x
- Piepho, H., Möhring, J., Melchinger, A. E., and Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161, 209–228. doi: 10.1007/s10681-007-9449-8
- Prakash, A. (2018). Cassava Market Developments and Outlook. In *Food Outlook-Biannual Report on Global Food Markets*. Rome: Food and Agriculture Organization of United Nations. doi: 10.1044/leader.PPL.19102014.18
- R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rabbi, Y. I., Kayondo, S. I., Bauchet, G., Yusuf, M., Aghogho, I. C., Ogunpaimo, K., et al. (2020). Genome-wide association analysis reveals new insights into the genetic architecture of defensive, agro-morphological and quality-related traits in cassava. *Plant Mol. Biol.* doi: 10.1007/s11103-020-01038-3 [Epub ahead of print].
- Rutkoski, J. E. (2019). Estimation of realized rates of genetic gain and indicators for breeding program assessment. *Crop Sci.* 59, 981–993. doi: 10.2135/cropsci2018.09.0537
- Shirima, R. R., Maeda, D. G., Kanju, E. E., Tumwegamire, S., Ceasar, G., Mushi, E., et al. (2019). Assessing the degeneration of cassava under high-virus inoculum conditions in coastal tanzania. *Plant Dis.* 103, 2652–2664. doi: 10.1094/PDIS-05-18-0750-RE
- Shittu, T. A., Alimi, B. A., Wahab, B., Sanni, L. O., and Abass, A. B. (2016). “Cassava flour and starch: processing technology and utilization,” in *Tropical Roots and Tubers: Production, Processing and Technology*, eds H. K. Sharma, N. Y. Njintang, R. S. Singhal, and P. Kaushal 415–450. doi: 10.1002/9781118992739.ch10a
- Spilke, J., Piepho, H., and Xiyuan, H. (2005). A simulation study on tests of hypotheses and confidence intervals for fixed effects in mixed models for blocked experiments with missing data. *J. Agric. Biol. Environ. Statist.* 10, 374–389. doi: 10.1198/108571105X58199
- Ssemakula, G. N., Baguma, Y., Van der Grif, R., Otim-Nape, G. W., and Orone, J. (2000). Improvement of local cassava germplasm in Uganda. *Uganda J. Agric. Sci.* 5, 8–11.
- Storey, H. H., and Nichols, R. F. W. (1938). Studies of the mosaic disease of cassava. *Ann. Appl. Biol.* 25, 790–806. doi: 10.1111/j.1744-7348.1938.tb02354.x
- Tumuhimbi, R., Shanahan, P., and Melis, R. (2014). Combining ability analysis of storage root yield and related traits in cassava at the seedling evaluation stage of breeding. *J. Crop Improvement* 28, 530–546. doi: 10.1080/15427528.2014.923798
- Wagaba, H., Beyene, G., Cynthia, T., Alicai, T., Fauquet, C. M., and Taylor, N. J. (2013). Efficient transmission of cassava brown streak disease viral pathogens by chip bud grafting. *BMC Res. Notes* 6:516.
- Wolfe, M. D., Rabbi, I. Y., Egesi, C., Hamblin, M., Kawuki, R., Kulakow, P., et al. (2016). Genome-wide association and prediction reveals genetic architecture of cassava mosaic disease resistance and prospects for rapid genetic improvement. *Plant Geno.* 9, 1–13. doi: 10.3835/plantgenome2015.11.0118

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 Manze, Rubaihayo, Ozimati, Gibson, Esuma, Bua, Alicai, Omongo and Kawuki. 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.



Fine Mapping of the Leaf Rust Resistance Gene *Lr65* in Spelt Wheat ‘Altgold’

Qiang Zhang, Wenxin Wei, Xiangxi Zuansun, Shengnan Zhang, Chen Wang, Nannan Liu, Lina Qiu, Weidong Wang, Weilong Guo, Jun Ma, Huiru Peng, Zhaorong Hu, Qixin Sun and Chaojie Xie*

State Key Laboratory for Agrobiotechnology, Key Laboratory of Crop Heterosis and Utilization (MOE), Key Laboratory of Crop Genetic Improvement, College of Agronomy and Biotechnology, China Agricultural University, Beijing, China

OPEN ACCESS

Edited by:

Harsh Raman,
New South Wales Department
of Primary Industries, Australia

Reviewed by:

Parveen Chhuneja,
Punjab Agricultural University, India
Elisabetta Mazzucotelli,
Council for Agricultural
and Economics Research (CREA),
Italy

*Correspondence:

Chaojie Xie
xiecj127@126.com

Specialty section:

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

Received: 11 February 2021

Accepted: 04 May 2021

Published: 28 June 2021

Citation:

Zhang Q, Wei W, Zuansun X,
Zhang S, Wang C, Liu N, Qiu L,
Wang W, Guo W, Ma J, Peng H,
Hu Z, Sun Q and Xie C (2021) Fine
Mapping of the Leaf Rust Resistance
Gene *Lr65* in Spelt Wheat ‘Altgold’.
Front. Plant Sci. 12:666921.
doi: 10.3389/fpls.2021.666921

Wheat leaf rust (also known as brown rust), caused by the fungal pathogen *Puccinia triticina* Erikss. (Pt), is one by far the most troublesome wheat disease worldwide. The exploitation of resistance genes has long been considered as the most effective and sustainable method to control leaf rust in wheat production. Previously the leaf rust resistance gene *Lr65* has been mapped to the distal end of chromosome arm 2AS linked to molecular marker *Xbarc212*. In this study, *Lr65* was delimited to a 0.8 cM interval between flanking markers *Alt-64* and *AltID-11*, by employing two larger segregating populations obtained from crosses of the resistant parent Altgold Rotkorn (ARK) with the susceptible parents Xuezaio and Chinese Spring (CS), respectively. 24 individuals from 622 F₂ plants of crosses between ARK and CS were obtained that showed the recombination between *Lr65* gene and the flanking markers *Alt-64* and *AltID-11*. With the aid of the CS reference genome sequence (IWGSC RefSeq v1.0), one SSR marker was developed between the interval matched to the *Lr65*-flanking marker and a high-resolution genetic linkage map was constructed. The *Lr65* was finally located to a region corresponding to 60.11 Kb of the CS reference genome. The high-resolution genetic linkage map founded a solid foundation for the map-based cloning of *Lr65* and the co-segregating marker will facilitate the marker-assisted selection (MAS) of the target gene.

Keywords: Altgold Rotkorn, *Lr65*, leaf rust resistance, fine mapping, marker-assisted selection

INTRODUCTION

Virtually anywhere wheat is cultivated, its production is seriously constrained by fungal pathogens, and most significantly by single or multiple of the three species of rust (Hovmöller et al., 2010), i.e., leaf rust (*Puccinia triticina*); stem rust (*Puccinia graminis* f. sp. *tritici*); and stripe rust (*Puccinia striiformis* f. sp. *tritici*). Among these, leaf rust is considered potentially the most disruptive disease due to its more frequency and widespread occurrence in all wheat-growing locations of the world (Roelfs et al., 1992; Bolton et al., 2008; Huerta-Espino et al., 2011). Leaf rust can cause a 15% production reduction and heavy infection can lead to losses of up to 40% (Knott, 1989; McMullen et al., 2008). Over the past decades, outbreaks of rust diseases have occurred in various regions of China, resulting in severe wheat yield reduction (Chen et al., 2018). Although leaf rust can be controlled through foliar fungicide applications, the most effective and eco-friendly way to control

the disease is based on improved varieties containing resistance genes (Keller et al., 2008). However, one of the most frustrating issues in disease resistance breeding is the failure of resistance genes, due to the evolving nature of plant pathogens resulting in new virulent races that can cause disease in formerly resistant wheat varieties. Therefore, it is necessary to search for new diverse effective resistance genes that can be used in wheat breeding programs.

To date, more than 80 leaf rust resistance genes (*Lr*) have been identified (Singh et al., 2013; Qureshi et al., 2018; Kumar et al., 2021). Roughly half of these genes are from wild relatives of wheat, while the remainder are from cultivated wheat (Marais et al., 2005; Naik et al., 2015; Rani et al., 2020). Wild relatives of wheat provide a huge gene pool of agronomy utility, including genes for rust resistance (Narang et al., 2019). The D genome donor of wheat, *Aegilops tauschii*, has been a rich source of resistance genes (Gill et al., 2019). Leaf rust resistance genes *Lr21*, *Lr32*, and *Lr39* have been transferred from *Ae. tauschii* into bread wheat (Raupp et al., 2001; Huang et al., 2003; Thomas et al., 2010). Tetraploid wheat is another important origin of disease resistance (Singh et al., 2017). *Lr14a*, *Lr23*, and *Lr53* were derived from durum wheat or wild emmer wheat and were used in common wheat breeding (McIntosh et al., 1995; Marais et al., 2005).

Spelt wheat (*Triticum spelta*) is an ancient crop that has been cultivated since 5000 BC (Xie et al., 2015). It is still a minor crop used for bread and fodder in Europe and North America today (Campbell, 1997). Spelt wheat has the same AABBDD genome as common wheat and their hybrids are fertile, facilitating the transfer of desirable genes to common wheat. In addition to genetic variation in protein concentration (Gomez-Becerra et al., 2010), lipid and mineral nutrient contents (Ruibal-Mendieta et al., 2002; Zhao et al., 2009), spelt wheat also shows excellent resistance to wheat rusts. Examples are the wheat yellow rust resistance gene *Yr5* gene, derived from spelt and localized on the long arm of chromosome 2B (Sun et al., 2002; Yan et al., 2003), and the *Lr44* gene in the Spelt variety 7831, located chromosome 1B (Dyck and Sykes, 1994).

Molecular markers have been used extensively in wheat breeding, principally for genetic mapping, marker-assisted selection (MAS), and positional gene cloning (Jost et al., 2020). With the evolution of sequencing technology, marker development has shifted to the sequencing era (Paux et al., 2012). The release of the annotated genome sequence of Chinese Spring (CS) have greatly improved our understanding of the wheat genome and facilitated the efficiency of marker development in wheat (Li et al., 2020).

The spelt wheat Altgold Rotkorn (ARK), a Swiss variety (Pedigree: Oberkulmer/Sandmeier), was first released in 1952. Wang et al. (2010) identified a leaf rust resistance gene *LrAlt* in Altgold and localized it to the distal end of the short arm of chromosome 2A. Mohler et al. (2012) reported the characterization and mapping of the same leaf rust resistance gene *LrARK0* in ARK. Since *LrAlt* and *LrARK0* were from the same germplasm and located at the same position, they were designated as *Lr65* (Mohler et al., 2012). In this study, we performed fine mapping of *Lr65* gene by exploring the CS reference genome. Our analysis located *Lr65* gene to a 60.11

Kb region on the IWGSC Ref-Seq v1.0 and identified one most likely candidate gene for *Lr65* in Altgold by comparing genome resequencing data between resistant and susceptible parents. In addition, co-segregating molecular markers were developed for MAS of the target gene.

MATERIALS AND METHODS

Plant and Pathogen Materials

Altgold, a spelt wheat cultivar with high resistance to leaf rust, was crossed with two susceptible common wheat lines “Xuezhao” and “CS”, and two F₂ segregating populations (Xuezhao/Altgold and CS/Altgold) were constructed. These two populations were used for the genetic analysis and mapping of leaf rust resistance gene. In all experiments, a susceptible common wheat line of Xuezhao was used as a comparison to check for successful inoculation. The *P. tritricina* isolate PHT (provided by Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China) was used for the inoculation. PHT was avirulent on Altgold and virulent on Xuezhao and CS. The conidia were propagated in the greenhouse on the susceptible plants.

Plant Growth and Pathogen Infection

The parental plants of Altgold, Xuezhao and CS and F₂ populations were tested for leaf rust resistance at seedling stage. The inoculations were initiated when the first leaves were fully unfolded, by spraying 1% Tween-20 aqueous solution as surfactant and then brushing conidia from the susceptible seedlings with sporulating leaf rusts onto the seedlings to be tested. The inoculated seedlings were incubated in dark plastic-covered boxes for 48 h at 15°C and 100% relative humidity and then transferred to greenhouse. 10–14 days after inoculation, infection types (ITs) were scored on a scale of 0–4 (0 = hypersensitive flecks, 1 = small uredinia with necrosis, 2 = moderate size pustules with chlorosis, 3 = moderate-large size uredinia without necrosis or chlorosis, and 4 = large uredinia lacking necrosis or chlorosis) (Stakman et al., 1962). ITs 0–2 represent resistance and ITs 3–4 represent susceptibility.

DNA Extraction and Quantification

DNA was extracted from seedlings of the F₂ populations and parents Altgold (resistant parent) as well as Xuezhao and CS (susceptible parents) using the CTAB method (Maroof et al., 1994). DNA samples were quantified using a NanoDrop One spectrophotometer instrument (Nanodrop Technologies) and diluted to a concentration of 30 ng/μl.

Resequencing of Resistant Parent Altgold

To obtain genomic variations between Altgold and CS, we performed whole-genome resequencing of Altgold. Altgold's whole genome sequencing was performed using the Illumina HiSeq2500 sequencing platform for double-end sequencing. The library construction and sequencing were performed by Beijing Novogene company. The read length of the paired-end sequencing library was 150 bp, the raw sequencing data

were processed according to GATK's best practices workflow (Van der Auwera et al., 2013).

Molecular Marker Development

Lr65 gene had already been mapped distal to marker *Xbarc212* on chromosome arm 2AS (Wang et al., 2010). Simple sequence repeats (SSRs) were developed based on the CS reference genome sequence distal to the *Xbarc212* locus. Meanwhile, InDels with insert/deletion size > 3bp were selected from the target interval between Altgold re-sequencing and CS reference genome sequence alignment database for further marker design. InDel polymerase chain reaction (PCR) primers were designed using Primer3Plus¹, with amplicon sizes ranging from 100 to 500 bp. BatchPrimer3 v1.0² was used to develop SSR markers.

Polymerase Chain Reaction Amplification and Visualization

Polymerase chain reaction amplification was performed in a 10 µL reaction volume containing 6 µL of 2 × Tag PCR StarMix with loading dye, 35–120 ng/mL DNA 2 µL, 1 µL of primer (mix of forward and reverse primers, 2 mM) and 1 µL of ddH₂O. The thermal profile consists of an initial denaturation step at 94°C for 5 min, followed by 35 cycles of 94°C for 30 s (denaturation), 50–61°C (depending on the annealing temperature of the specific primer) for 30 s, 72°C for 30 s (primer extension), and a terminal extension at 72°C for 10 min, stored at 4°C. The PCR products were separated by 10% non-denaturing polyacrylamide gel electrophoresis (acrylamide: bisacrylamide = 39:1), and gels were visualized with silver nitrate staining (Bassam et al., 1991).

Linkage Analysis and Map Construction

A chi-square analysis was performed on the leaf rust test data to confirm the goodness of fit of the observed ratios from the F₂ populations to the theoretical expected values. The χ^2 analysis was executed in Microsoft Excel (version 2010) using the Bchitest³ function to calculate χ^2 and *p*-values. The polymorphic markers tested between resistant and susceptible parents were used to genotype 2144 F₂ plants. The phenotypic data of disease responses were used for linkage analysis in combination with PCR amplification results. The localization of markers and the target gene is fulfilled based on recombination between markers genotype data and resistance/susceptibility phenotype data. Genetic distances were calculated in centiMorgan (cM).

Physical Mapping and Gene Annotation

The sequences of the two closest flanking markers linked to *Lr65* were used as lookups for a searches of the IWGSC RefSeq v1.0 to define the physical interval covering *Lr65* locus on CS chromosome 2AS. The gene annotation for the target interval was retrieved from the IWGSC RefSeq v1.0 annotation³.

Genomic Comparison Among Multiple Wheat Varieties

The sequence information was obtained from the Triticeae Multi-omics Center⁴ to obtain sequence information of annotated genes in candidate intervals, and then using the wheat 10 + genome⁵ (Walkowiak et al., 2020) for sequence alignment between the genomes of 15 wheat varieties.

RESULTS

Genetic Analysis of the Leaf Rust Resistance Gene *Lr65* in Two Segregating Populations

At the seedling stage, the parental lines Xueza0 and CS demonstrated a clear susceptible response to the leaf rust isolate PHT with an infection type (IT) score of 3, while Altgold showed a high-level resistant response with an IT score of 0 (Figures 1A,B). The F₁ plants and F₂ populations of Xueza0/Altgold and CS/Altgold were examined for the responses to the inoculation of the Pt isolate PHT at the seedling stage as well, along with the parents. The F₁ plants showed the same approximate immune infection type as the resistant parent Altgold, indicating the complete dominance of the resistance (Figures 1A,B). Of the 1522 F₂ plants screened from the Xueza0/Altgold cross, 1130 were resistant and 392 susceptible, fitting the ratio of 3:1 ($\chi^2_{3:1} = 0.46$, *p* > 0.05). In the F₂ population derived from cross CS/Altgold, 454 plants were resistant and 168 susceptible ($\chi^2_{3:1} = 1.33$, *p* > 0.05). The segregation of these two populations confirm that the leaf rust resistance in Altgold is controlled by a single dominant gene (Table 1), which is most likely the gene *Lr65* (previously known as *LrAlt* or *LrARK0*) (Wang et al., 2010; Mohler et al., 2012).

Marker Discovery and Molecular Mapping

Since *Lr65* gene has been located on the terminus of the short arm of chromosome 2A and the closest marker to *Lr65* is *Xbarc212* (Figure 2A) (Wang et al., 2010; Mohler et al., 2012). To further increase the map resolution in the *Lr65* region, new markers were developed using various genomic resources. 88 SSR primers were developed based on CS chromosome 2AS reference genome sequence (RefSeq v1.0) and tested on the two parents (Altgold and Xueza0). Four polymorphic markers (*Alt-14*, *Alt-21*, *Alt-24*, and *Alt-64*) were identified (Table 2). A total of 1522 Xueza0/Altgold F₂ plants were genotyped with these four markers, and 47 plants were identified with recombination between the marker loci and the resistance gene. Linkage analysis indicated that the closest marker to *Lr65* was *Alt-64* with a genetic distance of 0.5 cM. All four markers were on the proximal side to *Lr65* and closer than *Xbarc212* (Figure 2B).

To obtain markers on the other side of *Lr65*, we compared the resequencing data of Altgold with the CS reference in the

¹<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>

²<https://probes.pw.usda.gov/batchprimer3/>

³http://urgi.versailles.inra.fr/jbrowseiwgsc/gmod_jbrowse/

⁴<http://202.194.139.32/getfasta/index.html>

⁵https://webblast.ipk-gatersleben.de/wheat_ten_genomes/

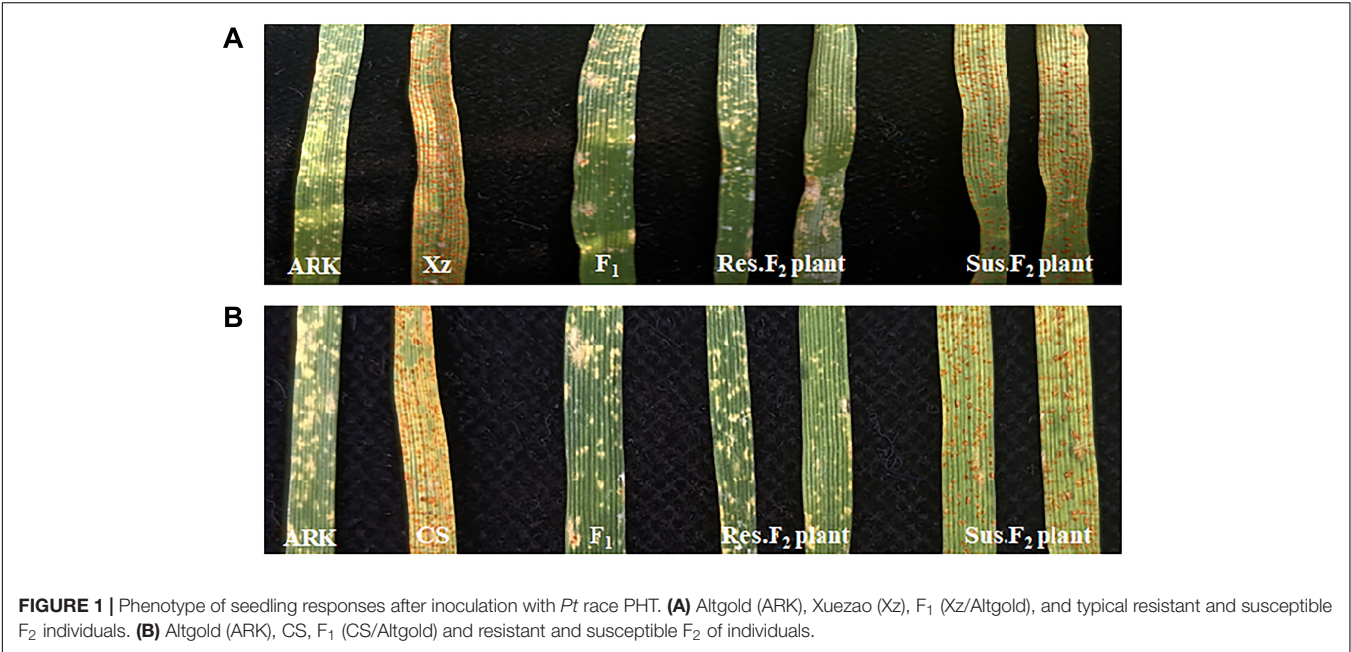
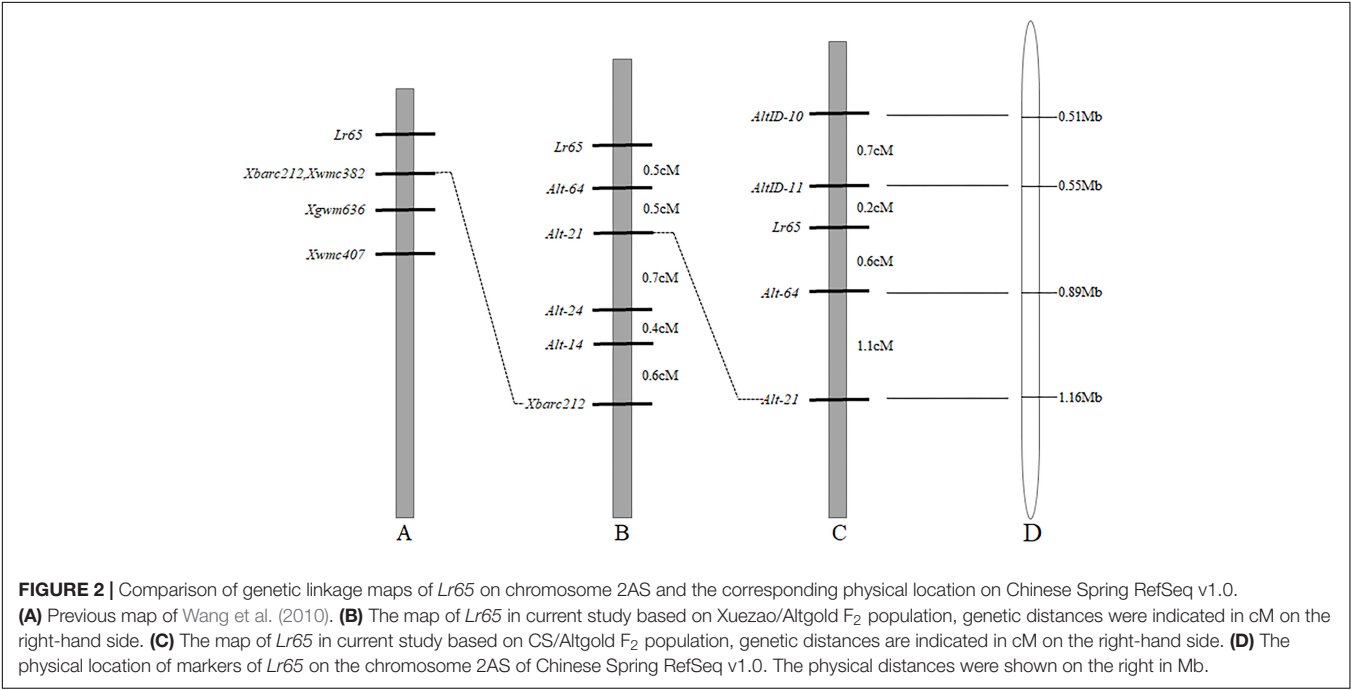


TABLE 1 | Segregation for leaf rust resistance in the Xuezao/Altgold and CS/Altgold F₂ population.

| Cross | Population | Number of seedling plants | | | χ^2 (3:1) |
|----------------|----------------|---------------------------|-------------|-------|---------------------------|
| | | Resistant | Susceptible | Total | |
| Xuezao/Altgold | F ₂ | 1130 | 392 | 1522 | $\chi^2 = 0.46, p = 0.49$ |
| CS/Altgold | F ₂ | 454 | 168 | 622 | $\chi^2 = 1.33, p = 0.24$ |



target region, which corresponds to the most distal 1.16 Mb interval of chromosome 2AS in CS RefSeq v1.0. Based on the Indel variations between the two parents, we designed eighteen Indel markers, two (*AltID-10* and *AltID-11*) of which were tested polymorphic between the parents (Table 2). These two Indel markers and previously developed SSR markers (*Alt-21* and *Alt-64*) were used to genotype 622 F₂ plants of the cross CS/Altgold. A genetic linkage map spanning 2.6 cM was constructed using these four markers (Figure 2C). In this map, *Lr65* gene is delimited to a genetic interval of 0.8 cM, flanked by markers *Alt-64* and *AltID-11*, with *AltID-11* 0.2cM distal to *Lr65* and *Alt-64* 0.6cM to *Lr65* on the proximal side.

When we matched the sequences of *Alt-64* and *AltID-11* with the genome sequence of CS (IWGSC v1.0), we found that the two markers were spanning an area of about 0.34 Mb (555551–891823) on CS chromosome 2AS (Figure 2D). Based on Altgold's re-sequencing data matching this 0.34 Mb interval, 11 SSR primers were designed and one more polymorphic marker *Alt-92* was found between Altgold and CS (Table 1). After tested among the 24 recombinants previously obtained by screening with the flanking markers *Alt-21*, *Alt-64* and *AltID-10*, *AltID-11*, two recombinants were identified between *Alt-92* and *Lr65*. These results showed that the *Lr65* locus was located between the markers *AltID-11* and *Alt-92* (Figure 3).

Physical Mapping and Gene Annotation of the *Lr65* Target Interval

In order to physically locate *Lr65*-linking markers, the sequences of all markers which were anchored in the high-resolution gene map were aligned to the CS reference genome sequence. The relative physical positions of these markers were generally consistent with the genetic linkage map (Figure 2). The closest flanking markers *AltID-11* and *Alt-92* of *Lr65* delimited a 60.11 Kb (555,551–615,668) interval in the CS Reference Genome (RefSeq v1.0). This region encompasses two annotated protein-coding genes, *TraesCS2A02G001400* and *TraesCS2A02G001500*, according to the IWGSC RefSeq v1.0 annotation⁶ (see text foot note 3) (Figure 3). The two annotated genes were put on NCBI⁷ to predict their protein structures, we found that *TraesCS2A02G001400* encodes a protein similar to that found in intracellular human pathogens with a conserved regions of internalin_A super family and *TraesCS2A02G001500* encodes a typical disease resistance protein (R protein) with a NB-ARC domain at the N-terminal end and three contiguous LRR at the C-terminal end (Supplementary Figure 1 and Table 3). One 3 bp Indel and one SNP were found in the coding sequence of *TraesCS2A02G001500* (Figure 4 and Supplementary Figure 3), indicating that these differences may lead to different protein functions, while there is no difference in sequence of *TraesCS2A02G001400* between Altgold and CS (Supplementary Figure 2). Therefore, *TraesCS2A02G001500* is most likely the candidate gene of *Lr65*.

⁶<http://ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>

⁷https://webblast.ipk-gatersleben.de/wheat_ten_genomes/

Comparison Among the Genomes of Multiple Wheat Cultivars

To validate the consistency of collinearity within candidate intervals in multiple wheat varieties, the wheat 10 + genome⁸ (Walkowiak et al., 2020) was used for comparison between genomes of additional wheat materials. Six wheat varieties (CS_RefSeq1.0, ArinaLrFor, Jagger, Julius, Norin61, Spelt) were identified in which both genes in the candidate interval were matched to chromosome arm 2AS, while three varieties were found matched to the same scaffold (Paragon_scaffold, Weebill_scaffold, and Cadenza_scaffold). The number of genes in these nine wheat varieties was consistent within the candidate interval, and the order of these two genes in these varieties was the same as in CS, with only one reversed (Cadenza_scaffold) (Figure 5). This indicates that the number of genes in the candidate region is uniform in multiple wheat varieties.

Development of the Diagnostic Marker of *Lr65*

Based on the 3-bp Indel in *TraesCS2A02G001500* between Altgold and CS, marker *1500-1* was developed and validated on Altgold, Xuezhao, and CS and the key recombinants (A25, A211, A321, and A523) (Table 2 and Figure 6). The test result indicated marker *1500-1* was co-segregating with *Lr65* gene (Figure 3).

In order to confirm the usefulness of this *Lr65* co-segregating marker in breeding, we tested marker *1500-1* on other 18 different Chinese wheat cultivars, we found that the PCR product size of marker *1500-1* in Altgold containing *Lr65* was unique and not detected in the other cultivars (Supplementary Figure 4); therefore, marker *1500-1* is diagnostic for selection of *Lr65* gene. Then we screened the marker *1500-1* in two other populations of F₁ progenies of crosses “Xuezhao/Altgold//Shiyou 20” and “Xuezhao/Altgold//Zhongmai 1062” and found that the marker was 100% associated with the leaf rust resistance (Figure 7 and Supplementary Figure 4). Since the resistant plants were the results of combining of *Lr65* with the susceptible alleles of Shiyou 20 and Zhongmai 1062, these plants all showed the heterozygous banding of marker *1500-1*.

DISCUSSION

In addition to *Lr65*, four wheat leaf rust resistance genes was located on the short arm of chromosome 2A, including *Lr17*, *Lr37*, and *Lr45* (McIntosh et al., 2008; Prasad et al., 2020). *Lr11* was previously located on chromosome 2AS (Soliman et al., 1964), but recent studies have shown that *Lr11* is located distal to chromosome 2DS (Darino et al., 2015). The gene *Lr17* has two resistance alleles, *Lr17a* and *Lr17b* (Dyck and Kerber, 1977; Singh et al., 2001). *Lr17a* was flanked by marker *Xgwm614* (distal) and *Xgwm407* (proximal), while marker *Xgwm636* was distal to *Xgwm614* (Bremenkamp-Barrett et al., 2008). *Lr65* (*LrAlt*) is mapped distal to *Xgwm636* (Wang et al., 2010). Gene *Lr37* is located within a fragment of *Ae. ventricosa* (Tausch) Cess. chromosome 2NS translocated to bread wheat chromosome 2AS, and genetic mapping analysis showed that the 2NS translocation

TABLE 2 | The primer sequences used in this study.

| Marker | Forward primer (5'–3') | Reverse primer (5'–3') | Marker type | Product size (bp) | Physical position (bp) | |
|------------------------|--------------------------|------------------------|-------------|-------------------|------------------------|---------|
| | | | | | Start | End |
| AltID-10 | CATCACTTTTGTCTCATCCA | CTATAACCCTGGCCCTTAATA | Indel | 153 | 517281 | 517434 |
| AltID-11 | AGAGGCTATGGATTGGAGTAG | CGCCATTAATGTCCATATCA | Indel | 249 | 555551 | 555800 |
| Alt-92 | GTCCTCTACAGTTCCATCC | GTGAAAACCATGTTGCAAAG | SSR | 206 | 615668 | 615874 |
| Alt-64 | AATCACATCACCCGACTCT | CGATTCTACCTTTCTGGACT | SSR | 173 | 891823 | 891996 |
| Alt-21 | GTAAAATAGAGGAGGGGTGAA | CATGTTAGAAGGGATAGAGAGG | SSR | 144 | 1166351 | 1166495 |
| Alt-24 | ACCCAATGCACCTGTACTCTAT | CTGGTGAATGGATGAAACA | SSR | 135 | 1227798 | 1227933 |
| Alt-14 | GCGAACAGAAAGAAAGAAAG | CCTAGACAGCACACATCTTGTA | SSR | 152 | 1309861 | 1310013 |
| 1500-1 | ATTCCATTGCCGTCTATCTT | GCACCTCCTTTTGTGTTG | Indel | 108 | 583323 | 583431 |
| Xbarc-212 ^a | GGCAACTGGAGTGATATAATACCG | CAGGAAGGGAGGAGAACAGAGG | SSR | 185 | 1582751 | 1582936 |

^aThe marker Xbarc212 was used in a previous study (Wang et al., 2010).

TABLE 3 | Candidate genes in the most distal 60.11kb region of 2AS.

| No. | Gene ID | Start position of the gene ^a | Length of gene (bp) | Gene annotation | Conservative regions |
|-----|--------------------|---|---------------------|---|---------------------------|
| 1 | TraesCS2A02G001400 | 562911 | 2915 | found in the intracellular human pathogen | internalin_A super family |
| 2 | TraesCS2A02G001500 | 580888 | 4532 | Disease resistance protein | NB-ARC and LRR |

^aGene ID and positions were fetched from the URGI website (<https://urgi.versailles.inra.fr/>) and as per IWGSC gene annotation v1.0.

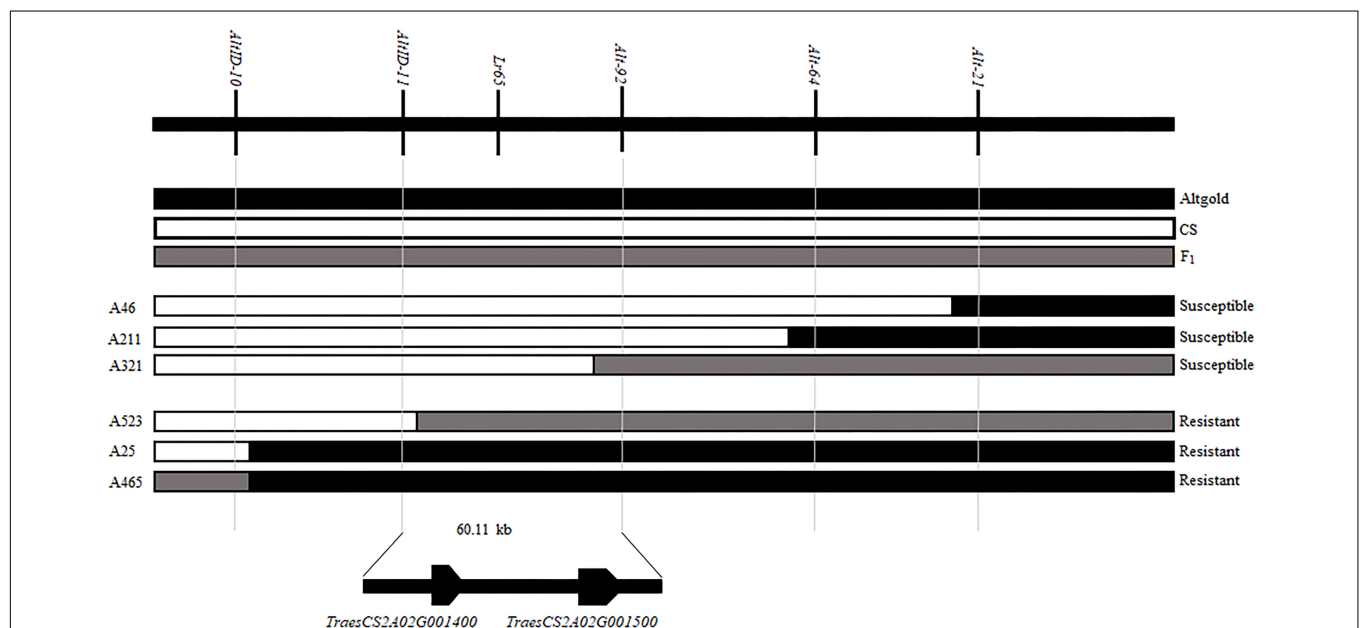


FIGURE 3 | Fine mapping of *Lr65* and two annotated genes in the target interval on the Chinese spring reference genome. The phenotypes and genotypes of six F_2 recombinants are displayed. The code and phenotype of each individual were put on the left and right sides, respectively. Black, white and gray blocks represent the genomic regions of Altgold, CS, and heterozygotes, respectively.

replaced about half of the short arm of chromosome 2A (Helguera et al., 2003). The gene *Lr45* is from rye chromosome 2R translocated to wheat chromosome 2A (Zhang et al., 2006). According to the above information, we conclude that *Lr65* is a unique leaf rust resistance gene.

Previously *Lr65* was mapped distal to the closest marker *Xbarc212* on wheat chromosome 2AS (Wang et al., 2010;

Mohler et al., 2012). In this study, using two large F_2 segregating populations of crosses Xueza0/Altgold and CS/Altgold, we fine mapped *Lr65* and narrow down it between markers AltID-11 and Alt-92, corresponding to the 60.11 Kb (555,551–615,668) interval according to the CS Reference Genome.

The gene fine mapping involved developing more polymorphic markers covering the genetic interval of the

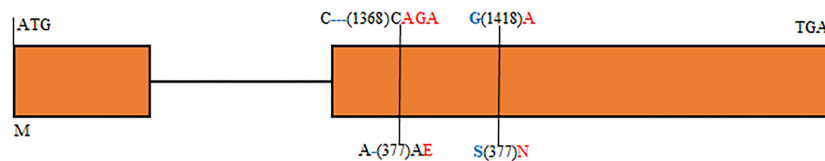


FIGURE 4 | Structure of the annotated gene *TraesCS2A02G001500* displaying nucleotide and amino acid sequence polymorphisms between the resistant and susceptible parents. Introns and exons are indicated by lines and orange boxes, respectively. Blue and red color characters indicate alleles of resistant and susceptible parents, respectively. Numbers in parentheses represent the positions of nucleotide and amino acid sequences relative to ATG and M. – Indicates a sequence deletion.

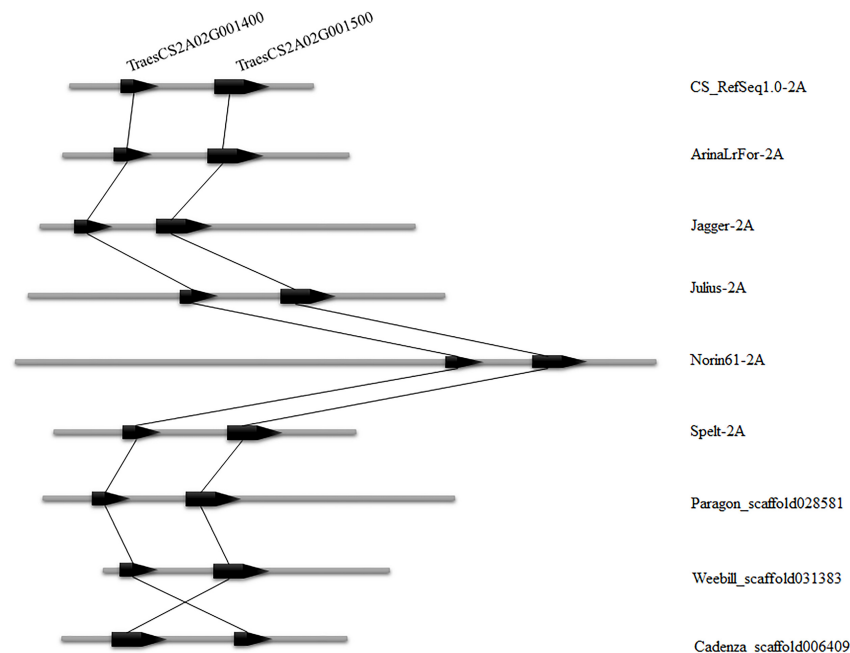


FIGURE 5 | Candidate interval of two genes *TraesCS2A02G001400* and *TraesCS2A02G001500* compared among nine wheat varieties. The left side indicates the position of the gene on the chromosome and the right side is the wheat variety name.

target gene. With the release of whole genome Reference sequence of CS, development of polymorphic markers associated with a target gene is becoming easier. The process of fine mapping of *Lr65* illustrates the effectiveness of the reference genome information and the resequencing data of the specific parental lines for the guided development of markers to target genes. Our work also demonstrate the advantage of using different crosses in the genetic mapping. Even though additional closer markers to *Lr65* were found using the Xueza0/Altgold F₂ population, all were on one side to the target gene. When we changed to the CS/Altgold population, the target gene was successfully delimited by flanking markers and narrow down to a shorter interval.

In the 60.11-Kb interval that contains *Lr65* locus on CS 2AS, there are two protein-coding genes annotated, *TraesCS2A02G001400* and *TraesCS2A02G001500*, according to the IWGSC RefSeq v1.0 annotation (see text foot note 3). Sequence analysis showed no difference in *TraesCS2A02G001400* between the resistant and susceptible parents (Altgold, CS and

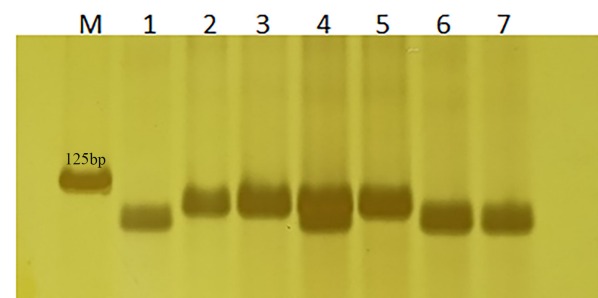


FIGURE 6 | Verification of *Lr65*'s diagnostic marker 1500-1. Nos. 1–3 are Altgold, Xueza0 and CS. Nos. 4–7 are recombinants A523, A25, A321, and A211, respectively. M is marker.

Xueza0). However, we found two sequence variations (one 3-bp Indel and one SNP) between Altgold and CS in the coding region of *TraesCS2A02G001500*. One marker was developed

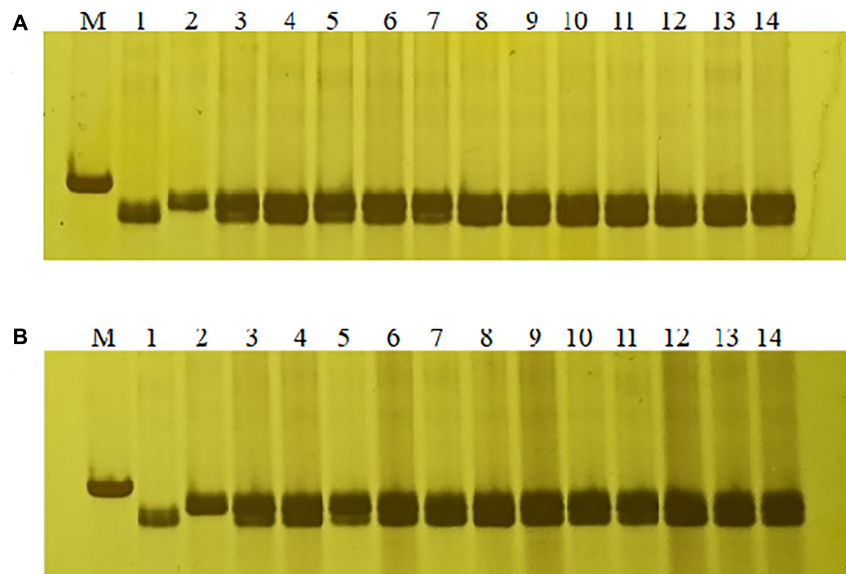


FIGURE 7 | Marker-assisted selection of *Lr65* with diagnostic marker 1500-1. **(A)** Selection of *Lr65* in “Xuexao/Altgold//Shiyou 20” population. 1: a resistant progeny of Xuexao/Altgold, 2: Shiyou20, 3–14: the resistant plants among the progenies. **(B)** Selection of *Lr65* in “Xuexao/Altgold//Zhongmai 1062” population. 1: a resistant progeny of Xuexao/Altgold, 2: Zhongmai1062, 3–14: the resistant plants among the progenies. M is marker.

to tag the 3-bp Indel variation between the parents and found to be co-segregating with *Lr65*. *TraesCS2A02G001500* was predicted to encode a protein with nucleotide binding sites and multiple leucine-rich repeats (NBS-LRR), the typical structures of disease resistance genes (R genes). Many cloned wheat rusts resistance genes are found to encode NBS-LRR proteins, including leaf rust resistance genes (*Lr1*, *Lr10*, *Lr21*, and *Lr22*) (Feuillet et al., 2003; Huang et al., 2003; Hiebert et al., 2007; Qiu et al., 2007), stripe rust resistance genes (*Yr5* and *Yr10*) (McGrann et al., 2014; Yuan et al., 2018), and stem rust resistance genes (*Sr22*, *Sr33*, *Sr35*, *Sr45*, and *Sr50*) (Saintenac et al., 2013; Periyannan et al., 2014; Casey et al., 2016; Saur et al., 2019; Md Hatta et al., 2020). Our results suggest that *TraesCS2A02G001500* might be the candidate gene of *Lr65*. The works to verify the disease resistance function of the Altgold allele of *TraesCS2A02G001500* are underway. However, there is still a chance that the sequence corresponding to *Lr65* is absent in CS genomic sequence. If so, we need to construct a genomic library of Altgold and to clone *Lr65* by physical mapping of contigs. The closest flanking and co-segregating markers developed in our present study will greatly aid the map-based cloning of *Lr65*.

Spelt is genetically distant from common wheat and with a high degree of genetic variation unexploited (Würschum et al., 2017; Akel et al., 2018). *Lr65* was first identified in spelt wheat and not being widely used in common wheat breeding. In addition to the resistance to the Chinese isolate PHT as in this study, *Lr65* was resistant to many Australia and Germany *P. triticulturae* isolates (Mohler et al., 2012). Utilization of *Lr65* will help to diversify the resistance genes in common wheat breeding and help to protect wheat production. However, due to the evolution

of new virulent pathogen isolates, major disease resistance genes are prone to lose their effectiveness when deployed alone. Mohler et al. (2012) had reported the existence of virulent pathotypes for *Lr65*. The *Lr65* gene was recommended to be used in combination with other resistance genes for the protection against leaf rust. The co-segregating marker we developed in present study would be helpful to pyramid *Lr65* with other resistance genes.

DATA AVAILABILITY STATEMENT

The datasets for this study can be found in the EVA repository (a preview of the data can be viewed here: <https://wwwdev.ebi.ac.uk/eva/?eva-study=PRJEB45547>). Project: PRJEB45547 and Analyses: ERZ2470822. SS ID: ss7173984454 and ss7173984455.

AUTHOR CONTRIBUTIONS

CX designed the research. QZ and CX conducted the research. CX, QZ, and WXW prepared the samples. QZ, XZ, SZ, CW, and NL analyzed the data. QZ wrote the draft. NL, LQ, CX, WG, JM, HP, ZH, and QS made the revision of the manuscript. All authors read and approved the manuscript.

FUNDING

This work was financially supported by the National Key Research and Development Project of China (2016YFD0101802) and the National Natural Science Foundation of China (31671676 and 31872865).

ACKNOWLEDGMENTS

We are grateful to Yinghui Li for assistance in revising the manuscript. We would like to thank Zihao Wang for the bioinformatics analyses.

REFERENCES

- Akel, W., Thorwarth, P., Mirdita, V., Weissman, E. A., Liu, G., Würschum, T., et al. (2018). Can spelt wheat be used as heterotic group for hybrid wheat breeding? *Theor. Appl. Genet.* 131, 973–984. doi: 10.1007/s00122-018-3052-3
- Bassam, B. J., Caetano-Anollés, G., and Gresshoff, P. M. (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal. Biochem.* 196, 80–83. doi: 10.1016/0003-2697(91)90120-1
- Bolton, M. D., Kolmer, J. A., and Garvin, D. F. (2008). Wheat leaf rust caused by *Puccinia triticina*. *Mol. Plant Pathol.* 9, 563–575. doi: 10.1111/j.13643703.2008.00487.x
- Bremenkamp-Barrett, B., Faris, J. D., and Fellers, J. P. (2008). Molecular mapping of the leaf rust resistance gene Lr17a in wheat. *Crop Sci.* 48, 1124–1128.
- Campbell, K. G. (1997). Spelt: agronomy, genetics, and breeding. *Plant Breed. Rev.* 15, 187–214. doi: 10.1002/9780470650097.ch6
- Casey, L. W., Lavrencic, P., Bentham, A. R., Cesari, S., Ericsson, D. J., Croll, T., et al. (2016). The CC domain structure from the wheat stem rust resistance protein Sr33 challenges paradigms for dimerization in plant NLR proteins. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12856–12861. doi: 10.1073/pnas.1609922113
- Chen, D., Shi, Y., Huang, W., Zhang, J., and Wu, K. (2018). Mapping wheat rust based on high spatial resolution satellite imagery. *Comput. Electron. Agric.* 152, 109–116. doi: 10.1016/j.compag.2018.07.002
- Darino, M. A., Dieguez, M. J., Singh, D., Ingala, L. R., Pergolesi, M. F., and Park, R. F. (2015). Detection and location of Lr11 and other leaf rust resistance genes in the durably resistant wheat cultivar Buck Poncho. *Euphytica* 206, 135–147.
- Dyck, P. L., and Kerber, E. R. (1977). Inheritance of leaf rust resistance in wheat cultivars Rafaela and EAP 26127 and chromosome location of gene Lr17. *Can. J. Genet. Cytol.* 19, 355–358.
- Dyck, P. L., and Sykes, E. E. (1994). Genetics of leaf-rust resistance in three spelt wheats. *Can. J. Plant Sci.* 74, 231–233. doi: 10.4141/cjps94-047
- Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A., and Keller, B. (2003). Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (*Triticum aestivum* L.) genome. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15253–15258. doi: 10.1073/pnas.2435133100
- Gill, H. S., Li, C., Sidhu, J. S., Liu, W., Wilson, D., Bai, G., et al. (2019). Fine mapping of the wheat leaf rust resistance gene Lr42. *Int. J. Mol. Sci.* 20:2445. doi: 10.3390/ijms20102445
- Gomez-Becerra, H. F., Erdem, H., Yazici, A., Tutus, Y., Torun, B., Ozturk, L., et al. (2010). Grain concentrations of protein and mineral nutrients in a large collection of spelt wheat grown under different environments. *J. Cereal Sci.* 52, 342–349. doi: 10.1016/j.jcs.2010.05.003
- Helguera, M., Khan, I. A., Kolmer, J., Lijavetzky, D., Zhong-Qi, L., and Dubcovsky, J. (2003). PCR assays for the Lr37-Yr17-Sr38 cluster of rust resistance genes and their use to develop isogenic hard red spring wheat lines. *Crop Sci.* 43, 1839–1847.
- Hiebert, C. W., Thomas, J. B., Somers, D. J., McCallum, B. D., and Fox, S. L. (2007). Microsatellite mapping of adult-plant leaf rust resistance gene Lr22a in wheat. *Theor. Appl. Genet.* 115, 877–884. doi: 10.1007/s00122-007-0604-3
- Hovmöller, M. S., Walter, S., and Justesen, A. F. (2010). Escalating threat of wheat rusts. *Science* 369–369. doi: 10.1126/science.1194925
- Huang, L., Brooks, S. A., Li, W., Fellers, J. P., Trick, H. N., and Gill, B. S. (2003). Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. *Genetics* 164, 655–664. doi: 10.1017/S0016672303006207
- Huerta-Espino, J., Singh, R. P., German, S., McCallum, B. D., Park, R. F., Chen, W. Q., et al. (2011). Global status of wheat leaf rust caused by *Puccinia triticina*. *Euphytica* 179, 143–160. doi: 10.1007/s10681-011-0361-x
- Jost, M., Singh, D., Lagudah, E., Park, R. F., and Dracatos, P. (2020). Fine mapping of leaf rust resistance gene Rph13 from wild barley. *Theor. Appl. Genet.* 133, 1887–1895. doi: 10.1007/s00122-020-03564-6
- Keller, B., Krattinger, S. G., Yahiaoui, N., Brunner, S., Kaur, N., Cloutre, C., et al. (2008). “Molecular analysis of fungal disease resistance in wheat,” in *Proceedings of the 11th international wheat genetics symposium*, vol. 1, eds R. Appels, R. Eastwood, E. Lagudah, P. Langridge, and M. Mackay, (Sydney, NSW: Sydney University Press Lynne McIntyre and Peter Sharp).
- Knott, D. R. (ed.) (1989). “The wheat rust pathogens,” in *The Wheat Rusts — Breeding for Resistance. Monographs on Theoretical and Applied Genetics*, vol. 12. (Berlin: Springer).
- Kumar, S., Bhardwaj, S., Gangwar, O., Sharma, A., Qureshi, N., Kumaran, V., et al. (2021). Lr80: a new and widely effective source of leaf rust resistance of wheat for enhancing diversity of resistance among modern cultivars. *Theor. Appl. Genet.* 134, 849–858. doi: 10.1007/s00122-020-03735-5
- Li, J., Yang, J., Li, Y., and Ma, L. (2020). Current strategies and advances in wheat biology. *Crop J.* 8, 879–891. doi: 10.1016/j.cj.2020.03.004
- Marais, G. F., Pretorius, Z. A., Wellings, C. R., McCallum, B., and Marais, A. S. (2005). Leaf rust and stripe rust resistance genes transferred to common wheat from *Triticum dicoccoides*. *Euphytica* 143, 115–123. doi: 10.1007/s10681-005-2911-6
- Maroof, M. S., Biyashev, R. M., Yang, G. P., Zhang, Q., and Allard, R. W. (1994). Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 91, 5466–5470. doi: 10.1073/pnas.91.12.5466
- McGrann, G. R., Smith, P. H., Burt, C., Mateos, G. R., Chama, T. N., MacCormack, R., et al. (2014). Genomic and genetic analysis of the wheat race-specific yellow rust resistance gene Yr5. *J. Plant Sci. Mol. Breed.* 3:2. doi: 10.7243/2050-2389-3-2
- McIntosh, R. A., Wellings, C. R., and Park, R. F. (1995). *Wheat Rusts: An Atlas of Resistance Genes*. Clayton, VIC: CSIRO publishing.
- McIntosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, W. J., Morris, C. F., and Sommers, D. J. (2008). “Catalogue of gene symbols for wheat: 2008,” in *Proceedings of the 11th International Wheat Genetics*, eds R. Appels, R. Eastwood, E. Lagudah, P. Langridge, M. Mackay, L. McIntyre, et al. (Sydney, NSW: Sydney University Press).
- McMullen, M., Markell, S. G., and Rasmussen, J. B. (2008). *Rust Diseases of Wheat in North Dakota*. Fargo: North Dakota State University Extension Bulletin, 1361.
- Md Hatta, M. A., Ghosh, S., Athiyannan, N., Richardson, T., Steuernagel, B., Yu, G., et al. (2020). Extensive genetic variation at the Sr22 wheat stem rust resistance gene locus in the grasses revealed through evolutionary genomics and functional analyses. *Mol. Plant Microbe Interact.* 33, 1286–1298. doi: 10.1094/MPMI-01-20-0018-R
- Mohler, V., Singh, D., Singrün, C., and Park, R. F. (2012). Characterization and mapping of Lr65 in spelt wheat ‘Altgold Rotkorn’. *Plant Breed.* 131, 252–257. doi: 10.1111/j.1439-0523.2011.01934.x
- Naik, B. K., Sharma, J. B., Sivasamy, M., Prabhu, K. V., Tomar, R. S., and Tomar, S. M. S. (2015). Molecular mapping and validation of the microsatellite markers linked to the Secale cereale-derived leaf rust resistance gene Lr45 in wheat. *Mol. Breed.* 35:61. doi: 10.1007/s11032-015-0234-4
- Narang, D., Kaur, S., Steuernagel, B., Ghosh, S., Dhillon, R., Bansal, M., et al. (2019). Fine mapping of Aegilops peregrina co-segregating leaf and stripe rust resistance genes to distal-most end of 5DS. *Theor. Appl. Genet.* 132, 1473–1485. doi: 10.1007/s00122-019-03293-5
- Paux, E., Sourdille, P., Mackay, I., and Feuillet, C. (2012). Sequence-based marker development in wheat: advances and applications to breeding. *Biotechnol. Adv.* 30, 1071–1088. doi: 10.1016/j.biotechadv.2011.09.015
- Periyannan, S., Bansal, U., Bariana, H., Deal, K., Luo, M. C., Dvorak, J., et al. (2014). Identification of a robust molecular marker for the detection of the stem rust resistance gene Sr45 in common wheat. *Theor. Appl. Genet.* 127, 947–955. doi: 10.1007/s00122-014-2270-6

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.666921/full#supplementary-material>

- Prasad, P., Savadi, S., Bhardwaj, S. C., and Gupta, P. K. (2020). The progress of leaf rust research in wheat. *Fungal Biol.* 124, 537–550.
- Qiu, J. W., Schürch, A. C., Yahiaoui, N., Dong, L. L., Fan, H. J., Zhang, Z. J., et al. (2007). Physical mapping and identification of a candidate for the leaf rust resistance gene Lr1 of wheat. *Theor. Appl. Genet.* 115, 159–168. doi: 10.1007/s00122-007-0551-z
- Qureshi, N., Bariana, H., Kumran, V. V., Muruga, S., Forrest, K. L., Hayden, M. J., et al. (2018). A new leaf rust resistance gene Lr79 mapped in chromosome 3BL from the durum wheat landrace Aus26582. *Theor. Appl. Genet.* 131, 1091–1098. doi: 10.1007/s00122-020-03625-w
- Rani, K., Raghu, B. R., Jha, S. K., Agarwal, P., Mallick, N., Niranjana, M., et al. (2020). A novel leaf rust resistance gene introgressed from *Aegilops markgrafii* maps on chromosome arm 2AS of wheat. *Theor. Appl. Genet.* 133, 2685–2694. doi: 10.1007/s00122-020-03625-w
- Raupp, W. J., Brown-Guedira, G. L., and Gill, B. S. (2001). Cytogenetic and molecular mapping of the leaf rust resistance gene Lr39 in wheat. *Theor. Appl. Genet.* 102, 347–352. doi: 10.1007/s001220051652
- Roelfs, A. P., Singh, R. P., and Saari, E. E. (1992). *Rust Diseases of Wheat: Concepts and Methods of Disease Management*. Mexico: International Maize and Wheat Improvement Center (CIMMYT), 1–18
- Ruibal-Mendieta, N. L., Delacroix, D. L., and Meurens, M. (2002). A comparative analysis of free, bound and total lipid content on spelt and winter wheat wholemeal. *J. Cereal Sci.* 35, 337–342. doi: 10.1006/jcrs.2001.0434
- Saintenac, C., Zhang, W., Salcedo, A., Rouse, M. N., Trick, H. N., Akhunov, E., et al. (2013). Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science* 341, 783–786. doi: 10.1126/science.1239022
- Saur, I. M., Bauer, S., Lu, X., and Schulze-Lefert, P. (2019). A cell death assay in barley and wheat protoplasts for identification and validation of matching pathogen AVR effector and plant NLR immune receptors. *Plant Methods* 15:118. doi: 10.1186/s13007-019-0502-0
- Singh, A. K., Sharma, J. B., Singh, P. K., Singh, A., and Mallick, N. (2017). Genetics and mapping of a new leaf rust resistance gene in *Triticum aestivum* L. × *Triticum timopheevii* Zhuk. derivative ‘Selection G12’. *J. Genet.* 96, 291–297. doi: 10.1007/s12041-017-0760-4
- Singh, D., Mohler, V., and Park, R. F. (2013). Discovery, characterisation and mapping of wheat leaf rust resistance gene Lr71. *Euphytica* 190, 131–136. doi: 10.1007/s10681-012-0786-x
- Singh, D., Park, R. F., Bariana, H. S., and McIntosh, R. A. (2001). Cytogenetic studies in wheat XIX. Chromosome location and linkage studies of a gene for leaf rust resistance in the Australian cultivar ‘Harrier’. *Plant Breed.* 120, 7–12.
- Soliman, A. S., Heyne, E. G., and Johnston, C. O. (1964). Genetic analysis for leaf rust resistance in the eight differential varieties of wheat 1. *Crop Sci.* 4, 246–248.
- Stakman, E. C., Stewart, D. M., and Loegering, W. Q. (1962). *Identification of Physiologic Races of Puccinia Graminis var. tritici*. Washington, DC: USDA.
- Sun, Q., Wei, Y., Ni, Z., Xie, C., and Yang, T. (2002). Microsatellite marker for yellow rust resistance gene Yr5 in wheat introgressed from spelt wheat. *Plant Breed.* 121, 539–541. doi: 10.1046/j.1439-0523.2002.00754.x
- Thomas, J., Nilmalgoda, S., Hiebert, C., McCallum, B., Humphreys, G., and DePauw, R. (2010). Genetic markers and leaf rust resistance of the wheat gene Lr32. *Crop Sci.* 50, 2310–2317. doi: 10.2135/cropsci2010.02.0065
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* 43, 11–10.
- Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588, 277–283. doi: 10.1038/s41586-020-2961-x
- Wang, Y., Peng, H., Liu, G., Xie, C., Ni, Z., Yang, T., et al. (2010). Identification and molecular mapping of a leaf rust resistance gene in spelt wheat landrace Altgold. *Euphytica* 174, 371–375. doi: 10.1007/s10681-010-0134-y
- Würschum, T., Leiser, W. L., and Longin, C. F. H. (2017). Molecular genetic characterization and association mapping in spelt wheat. *Plant Breed.* 136, 214–223. doi: 10.1111/pbr.12462
- Xie, Q., Mayes, S., and Sparkes, D. L. (2015). Spelt as a genetic resource for yield component improvement in bread wheat. *Crop Sci.* 55, 2753–2765. doi: 10.2135/cropsci2014.12.0842
- Yan, G., Chen, X., Line, R., and Wellings, C. (2003). Resistance gene-analog polymorphism markers co-segregating with the Yr5 gene for resistance to wheat stripe rust. *Theor. Appl. Genet.* 106, 636–643. doi: 10.1007/s00122-002-1109-8
- Yuan, C., Wu, J., Yan, B., Hao, Q., Zhang, C., Lyu, B., et al. (2018). Remapping of the stripe rust resistance gene Yr10 in common wheat. *Theor. Appl. Genet.* 131, 1253–1262. doi: 10.1007/s00122-018-3075-9
- Zhang, N., Yang, W. X., Yan, H. F., Liu, D. Q., Dong, C. H. U., Meng, Q. F., et al. (2006). Molecular markers for leaf rust resistance gene Lr45 in wheat based on AFLP analysis. *Agric. Sci. China* 5, 938–943.
- Zhao, F. J., Su, Y. H., Dunham, S. J., Rakszegi, M., Bedo, Z., McGrath, S. P., et al. (2009). Variation in mineral micronutrient concentrations in grain of wheat lines of diverse origin. *J. Cereal Sci.* 49, 290–295. doi: 10.1016/j.jcs.2008.11.007

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 Zhang, Wei, Zuansun, Zhang, Wang, Liu, Qiu, Wang, Guo, Ma, Peng, Hu, Sun and Xie. 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.



Outlook of Cassava Brown Streak Disease Assessment: Perspectives of the Screening Methods of Breeders and Pathologists

Alfred A. Ozimati^{1*}, Williams Esuma¹, Titus Alicai¹, Jean-Luc Jannink², Chiedozi Egesi² and Robert Kawuki¹

¹ Root Crops Program, National Crops Resources Research Institute (NaCRRI), Kampala, Uganda, ² College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Francis Chuks Ogbonnaya,
Grains Research and Development
Corporation, Australia
Jiao Liu,
Chinese Academy of Tropical
Agricultural Sciences, China

*Correspondence:

Alfred A. Ozimati
ozimatialfred@gmail.com

Specialty section:

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

Received: 31 December 2020

Accepted: 31 May 2021

Published: 05 July 2021

Citation:

Ozimati AA, Esuma W, Alicai T,
Jannink J-L, Egesi C and Kawuki R
(2021) Outlook of Cassava Brown
Streak Disease Assessment:
Perspectives of the Screening
Methods of Breeders and
Pathologists.
Front. Plant Sci. 12:648436.
doi: 10.3389/fpls.2021.648436

Cassava production and productivity in Eastern, Central, and Southern Africa are ravaged by cassava brown streak disease (CBSD), causing yield losses of up to 100% when susceptible varieties are grown. Efforts to develop CBSD-resistant clones are underway. However, the methods for screening CBSD resistance currently vary between breeders and pathologists, with the limited empirical data to support their choices. In this study, we used the empirical CBSD foliar and root necrosis data from two breeding populations, termed cycle zero (C_0) and cycle one (C_1), to assess and compare the effectiveness of the CBSD screening methods of breeders vs. pathologists. On the one hand, the estimates of broad-sense heritability (H^2) for the CBSD root necrosis assessment of breeder ranged from 0.15 to 0.87, while for the assessment method of pathologists, H^2 varied from 0.00 to 0.71 in C_0 clones. On the other hand, the marker-based heritability estimates (h^2) for C_0 ranged from 0.00 to 0.70 for the assessment method of breeders and from 0.00 to 0.63 for the assessment method of pathologists. For cycle one (C_1) population, where both foliar and root necrosis data were analyzed for clones assessed at clonal evaluation trials (CETs) and advanced yield trials (AYTs), H^2 varied from 0.10 to 0.59 for the assessment method of breeders, while the H^2 values ranged from 0.09 to 0.35 for the CBSD computation method of pathologists. In general, higher correlations were recorded for foliar severity from the assessment method of breeders ($r = 0.4$, $p \leq 0.01$ for CBSD3s and $r = 0.37$, $p \leq 0.01$ for CBSD6s) in C_1 clones evaluated at both clonal and advanced breeding stages than from the approach of pathologists. Ranking of top 10 C_1 clones by their indexed best linear unbiased predictors (BLUPs) for CBSD foliar and root necrosis showed four overlapping clones between clonal and advanced selection stages for the method of breeders; meanwhile, only a clone featured in both clonal and advanced selection stages from the CBSD assessment method of pathologists. Overall, the CBSD assessment method of breeders was more effective than the assessment method of pathologists, and thus, it justifies its continued use in CBSD resistance breeding.

Keywords: cassava, resistance, breeder's, pathologist's, cassava brown streak disease

INTRODUCTION

The human population in the next 30 years is projected to increase by 25%, from the current world population of ~7.5 billion to 10 billion people. The highest rate of this growth is expected to arise from sub-Saharan Africa (SSA; Hickey et al., 2017). Consequently, there is an urgent need to match this rapid growth in the human population with a concomitant increase in food production. Cassava (*Manihot esculenta* Crantz), a climate-resilient food staple in SSA, is a suitable crop to meet the projected calorie demand since more than half of the global production is in Africa (FAOSTAT, 2019).

Unfortunately, the average on-farm yield of cassava in Africa is low, stagnating at 12 tons/ha compared with 20 tons/ha estimated for Asian and Latin American countries (Malik et al., 2020). The biotic factors, such as cassava brown streak disease (CBSD), cassava mosaic disease (CMD), cassava bacterial blight, and whitefly vector, are the key obstacles to optimal cassava production and productivity in Africa (Maruthi et al., 2005; Mware et al., 2009; Patil and Fauquet, 2009; Patil et al., 2015). In the case of East Africa, the CBSD is currently the most devastating constraint for cassava production, causing yield losses of up to 100% in highly susceptible varieties (Alicai et al., 2007; Legg et al., 2011; Hillocks and Maruthi, 2015). Typical cassava plants infected with CBSD present characteristic yellowing along the veins, compromising the photosynthetic capacity of leaves, brown streaks on stems, and corky necrosis in the edible root parenchyma, and rendering the roots unusable for food or feed (Hillocks, 2004; Patil et al., 2015; Hillocks et al., 2016).

The severity and incidence of foliar and root CBSD symptoms form the basis of CBSD resistance screening. Currently, a scale of 1–5 is used to independently assess CBSD severity on foliar and roots; these assessments are commonly performed at 3 (CBSD3s) and 6 (CBSD6s) months for foliar and at 12 (CBSDRs) months at harvest for root necrosis (Hillocks, 2004; Kaweesi et al., 2014; Okul et al., 2018). The scores for the foliar severity assessment are as follows: 1 = no symptom, 2 = slight foliar chlorotic leaf mottle with no stem lesions, 3 = foliar chlorotic leaf mottle and blotches with mild stem lesions, 4 = foliar chlorotic leaf mottle and blotches with well-pronounced stem lesions, but no dieback, and 5 = defoliation with stem lesions and dieback. The scores for the root necrosis assessment are as follows: 1 = no necrosis, 2 = mild necrotic lesions (1–10%), 3 = pronounced necrotic lesions (11–25%), 4 = severe necrotic lesion (26–50%), and 5 = very severe necrotic lesions (>50%).

Although the CBSD symptom expressions are common to both breeders and pathologists, there is an apparent discrepancy in the data processing for decision support. For example, pathologists compute plot scores by averaging all severity scores ≥ 2 , i.e., they exclude the CBSD severity scores of 1 when deriving plot mean for foliar and root symptoms (Ogwok et al., 2012; Odipio et al., 2014; Wagaba et al., 2017). On the other hand, breeders compute the averages of CBSD foliar and root severity using all the recorded observations, i.e., they do not exclude the CBSD scores of 1 (Kawuki et al., 2016, 2019; Okul et al., 2018). Essentially, the average values obtained from the CBSD

assessments of pathologists or breeders are the different traits used for decision support.

In our efforts to optimize the cassava breeding operations tailored toward increased genetic gains, there is a need to assess the precision and relationship between the CBSD assessment methods. A key metric used to assess trait reliability is heritability, which measures the ratio of genetic variance to phenotypic variance (broad-sense heritability) or the ratio of additive genetic variance to phenotypic variance (narrow-sense heritability) (Bernardo, 2003). Accordingly, the data sets presented in this study aimed at answering the following research questions: (a) What proportion of total genetic and additive genetic variances are captured by the CBSD assessment methods of breeders and pathologists? and (b) To what extent do the CBSD assessment methods of breeders and pathologists select and advance the same clones?

MATERIALS AND METHODS

Test Clones and CBSD Field Evaluations

The clones used in this study comprised genomic selection cycle zero (C_0) and cycle one (C_1) populations developed by the cassava breeding program of National Crops Resources Research Institute (NaCRRI). The data for C_0 clones presented in this study were collected from clonal evaluation trials (CETs), while C_1 clones were evaluated in both CETs and advanced yield trials (AYTs). The first set of CETs from C_0 , herein referred to as CETs-1, were evaluated at seven sites during first (April–May) and second (September–October) planting seasons in 2015. The first and second plantings generally depict the onset of rains. In Uganda, our first and second rains typically appear in February–March and September–October, respectively. The trial sites represent some of the key cassava production and consumption zones in Uganda. In these multilocal trials, a total of 155 C_0 clones from a genomic selection training population of 427 genotypes were evaluated (Ozimati et al., 2018). Each trial was established in an augmented design with five checks (i.e., UG110008, UG110014, UG110015, UG110016, and UG110017) and replicated—five to six times in single-row plots of 10 plants spaced at 1×1 m between and within rows.

On the other hand, the C_1 population presented in this study was generated from crosses made among 100 progenitors, a subset of the 155 C_0 clones. In 2015–2016, we started with a seedling evaluation of ~5,000 genotypes for C_1 , from which 735 clones were evaluated in CET (2016–2017), herein referred to as CETs-2 at two locations (i.e., Namulonge and Serere). The CETs-2 were also planted in an augmented design with three checks, namely, UG110015, UG110017, and UG110134 in single-row plots of 10 plants spaced at 1×1 m between and within rows. During harvest in August 2017, a subset of 50 C_1 clones were selected, based on the yield performance and response to CBSD as well as CMD from the CETs-2, and established in AYT at three locations (i.e., Arua, Serere, and Namulonge). At each location, the trials were established in randomized complete block design, with a plot size of 6×6 m, replicated twice. For all trials, the plots were separated by 2-m alleys.

Since the plant-based foliar CBSD data collected at 3 (CBSD3s) and 6 (CBSD6s) months after planting (MAP) were only available for C_1 clones assessed at CETs-2 and AYT, we derived the mean foliar CBSD values for the assessment methods of breeders and pathologists for this population. To compute the plot means for foliar CBSD severity for the two disease assessment methods, plant-based diseases scored on a scale of 1–5 were used. In this case, score 1 = no foliar symptom expressions, 2 = mild symptoms (1–10%), 3 = pronounced chlorotic mottle and mild stem lesions (11–25%), 4 = foliar chlorotic leaf mottle and blotches with pronounced stem lesions (26–50%), and 5 = defoliation with stem lesions and dieback (>50%) (Hillocks and Thresh, 2000).

At harvest, which coincided with 12 MAP for both C_0 and C_1 populations, all plants per plot were uprooted, and roots were also assessed individually for CBSD necrosis using the scale of 1–5, where 1 = no necrosis, 2 = mild necrotic lesions (1–10%), 3 = pronounced necrotic lesions (11–25%), 4 = severe necrotic lesions (26–50%) with mild root constrictions, and 5 = very severe necrotic lesions (>50%) with severe root constrictions (Hillocks and Thresh, 2000; Kaweesi et al., 2014). We further processed the root necrosis data to match the mean CBSD severity computation methods of breeders and plant pathologists, i.e., all root severity scores were averaged for the assessment method of breeders, while only the root severity scores ≥ 2 were averaged for the CBSD assessment method of pathologists.

Genotyping of the Clones

DNA was extracted from ~100 mg of fresh young leaves from each of the 155 C_0 clones. DNA extractions were performed using QIAGEN DNeasy, Texas, USA extraction kits and quantified using Picogreen® to ensure that the required concentrations for sequencing were obtained. Consequently, DNA samples were genotyped using the genotyping-by-sequencing method as described by Elshire et al. (2011). Removing the single nucleotide polymorphic (SNP) markers by filtering and imputation methods has been described in an earlier study (Hamblin and Rabbi, 2014; Wolfe et al., 2016, 2017). Ultimately, we had a total of 25,383 SNP markers, which were filtered at minor allele frequency (MAF) ≥ 0.01 for the estimation of SNP-based heritability for each of the C_0 clones.

Statistical Analyses

To estimate the broad-sense heritability for each CBSD assessment method, i.e., breeders vs. pathologists for C_0 clones, we fitted the linear mixed model for each trial using the *lme4* package for the R statistical computing software (R Development Core Team, 2008) as follows:

$$y_{ijk} = \mu + c_i + b_j + e_{ijk} \quad \text{Model 1}$$

where y_{ijk} was the response of i th clone from j th block in the k th plot, μ represented the fixed trial mean, b and c represented a vector of random block and clone effects, respectively, and e was the random residual term. The variance components to compute the broad-sense heritability (H^2) were extracted from the model described earlier. The plot-based broad-sense heritability

estimates for root necrosis for the two CBSD assessment methods across 14 CETs-1 (i.e., location–season combinations) were then computed as follows:

$$H^2 = \frac{\sigma_c^2}{(\sigma_c^2 + \sigma_b^2 + \sigma_e^2)}$$

where σ_c^2 was the clone variance, σ_b^2 was the variance due to blocks, and σ_e^2 was the model residual variance.

To obtain the genomic estimated breeding values and the additive genetic variance for the two methods from CETs-1, we fitted a single-step genomic best linear unbiased predictor (G-BLUP) model as follows:

$$y_{ijk} = \mu + w_i + g_j + e_{ijk} \quad \text{Model 2}$$

where y_{ijk} was the response of j th genotype in the i th block recorded for k th plot, μ and w were the fixed grand mean and block effects, respectively, g_j represented the random genotype effect, assuming $g_j \sim N(0, G\sigma_g^2)$ with σ_g^2 representing the variance due to genotypic effects while G represented the covariance structure among clones based on the marker data, and e was the random model residual effect, assumed to be normally distributed as $e_{ijk}^{IID} \sim N(0, \sigma_e^2)$ with σ_e^2 as the residual variance. We extracted the variance components from the G-BLUP model and estimated the narrow-sense heritability (h^2 SNP-heritability) using the formula as follows:

$$h^2 = \frac{\sigma_g^2}{(\sigma_g^2 + \sigma_e^2)}$$

where σ_g^2 was the additive genetic variance and σ_e^2 was the model residual variance.

Furthermore, we examined how many top 10 ranked clones at CETs-2 were featured among the best 10 clones at AYT for the two CBSD assessment methods from C_1 population. To do this, the data sets from each of the two trial stages (i.e., CETs-2 and AYT) were combined across sites, followed by fitting a multilocal linear mixed model as described below for each trial stage. For the CETs-2, we fitted a multilocal model described as follows:

$$y_{ijkl} = \mu + l_i + g_j + b/l_{k(i)} + gl_{ij} + e_{ijkl} \quad \text{Model 3}$$

where the grand mean μ and the main effect of the i th environment (l) were considered fixed, while the j th genotype (g), the k th block (b) nested within the i th environment (l), the interaction of the j th genotype (g) by i th environment (gl), and the residual term (e) were considered random. The variance components were extracted for the estimation of broad-sense heritability, using the formula described above for CETs-1.

Similarly, we fitted a multilocal linear mixed model for C_1 AYT, where the grand mean and location were considered fixed, while clones, replicates nested within trial, genotype-by-environment interactions, and residual terms were considered random. Accordingly, the variance components were extracted

to compute the plot-based broad-sense heritability estimates for foliar and root necrosis for the two CBSD assessment methods.

The raw phenotypic means and BLUP values for foliar CBSD severity as well as root necrosis of C_1 clones were extracted for both CETs-2 and AYT from the models fitted and used to compute Pearson's correlation coefficients for 50 C_1 clones that featured in both CETs-2 and AYT for each of the CBSD assessment methods. Furthermore, we computed selection index (SI) from BLUPs and raw phenotypic means of the three traits across sites, with the traits having equal economic weights as follows:

$$SI = -1(\text{CBSD3s}) + -1(\text{CBSD6s}) + -1(\text{CBSDRs})$$

where CBSD3s, CBSD6s, and CBSDRs were the CBSD severities assessed at 3, 6, and 12 MAP, respectively.

Finally, we used the indexed BLUP values of the three traits for the 50 clones that appeared at both CETs-2 and AYT for ranking the top 10 clones at each trial stage. The purpose of ranking was to compare the number of 10 top clones that overlapped at CETs-2 and AYT for each of the CBSD averaging methods.

RESULTS

Broad-Sense and SNP-Heritability Estimates

The broad-sense heritability (H^2) estimates for the CBSD root severity assessment method of breeders ranged from 0.15 in Arua 2015A trial to 0.87 in Namulonge 2015A trial (Table 1). On the other hand, H^2 estimates for the assessment method of pathologists ranged from 0.00 in Arua 2015A trial to 0.71 in Namulonge 2015A and B trials (Table 1). Meanwhile, the narrow-sense heritability (h^2) estimates, also referred to as SNP-based heritability, for the assessment method of breeders ranged from 0.00 in Arua 2015A trial to 0.72 in Namulonge 2015A trial (Table 1). Similarly, h^2 for the assessment method of pathologists varied from 0.00 in Arua 2015A trial to 0.63 in Serere 2015A trial. Overall, the average broad-sense and narrow-sense heritability estimates across trials were higher for the CBSD assessment method of breeders ($H^2 = 0.56$ and $h^2 = 0.36$) than for the CBSD assessment approach of pathologists ($H^2 = 0.49$ and $h^2 = 0.25$) (Table 1).

For C_1 population, the broad-sense heritability estimates for foliar and root necrosis from both CETs-2 and AYT are presented in Figure 1. We also observed higher H^2 values for the CBSD assessment method of breeders compared with the CBSD assessment method of pathologists for both CET and AYT evaluation stages. For example, at CET, H^2 at 3 months was 0.48 for the method of breeders and 0.38 for the method of pathologists. At 6 months, H^2 was 0.47 for the method of breeders and 0.21 for the computation of pathologists. Based on the root necrosis data at harvest, the broad-sense heritability values were 0.44 and 0.35 for the methods of breeders and pathologists, respectively. Similarly, the higher broad-sense heritability estimates of 0.42 and 0.56 were recorded for the combined data from AYT for the method of breeders compared

with the estimates of 0.41 and 0.09 recorded for the computations of pathologists for CBSD3s and CBSD6s, respectively (Figure 1).

Relationship Between BLUP Values of the 50 Clones Evaluated at CETs-2 and AYT for Mean CBSD Assessment Methods

In general, we recorded higher Pearson's correlation coefficients from the foliar CBSD assessment method of breeders than the approach of pathologists, using both BLUP estimates and raw phenotypic means across locations (Table 2). On the one hand, the highest correlation coefficient value ($r = 0.40$, $p \leq 0.01$) was observed for CBSD3s from the assessment method of breeders. On the other hand, low and statistically nonsignificant correlation coefficients were recorded for root necrosis and indexed trait values for both the CBSD assessment methods (Table 2). The correlation values for root necrosis and indexed trait values varied from 0.02 to 0.21. Overall, for all three disease traits and their indexed values, the CBSD computation method of breeders had higher correlation coefficients than the approach of the CBSD assessment of pathologists (Table 2).

Ranking of 50 Clones in CETs-2 and AYT Using Indexed BLUPs Values for the Two CBSD Averaging Methods

We ranked the 50 clones from C_1 , CETs-2, and AYT by their indexed BLUP values of CBSD3s, CBSD6s, and CBSDRs for the two CBSD assessment methods (Table 3). Based on ranking of the top 10 clones, four clones (i.e., UG15F190P001, UG15F170P507, UG15F079P011, and UG15F176P502) evaluated in CETs-2 and AYT overlapped among the top 10 ranked clones for the mean CBSD assessment method of breeders, whereas only one clone (UG15F190P001), overlapped between CETs-2 and AYT evaluated among the top 10 ranked clones (Table 3).

DISCUSSION

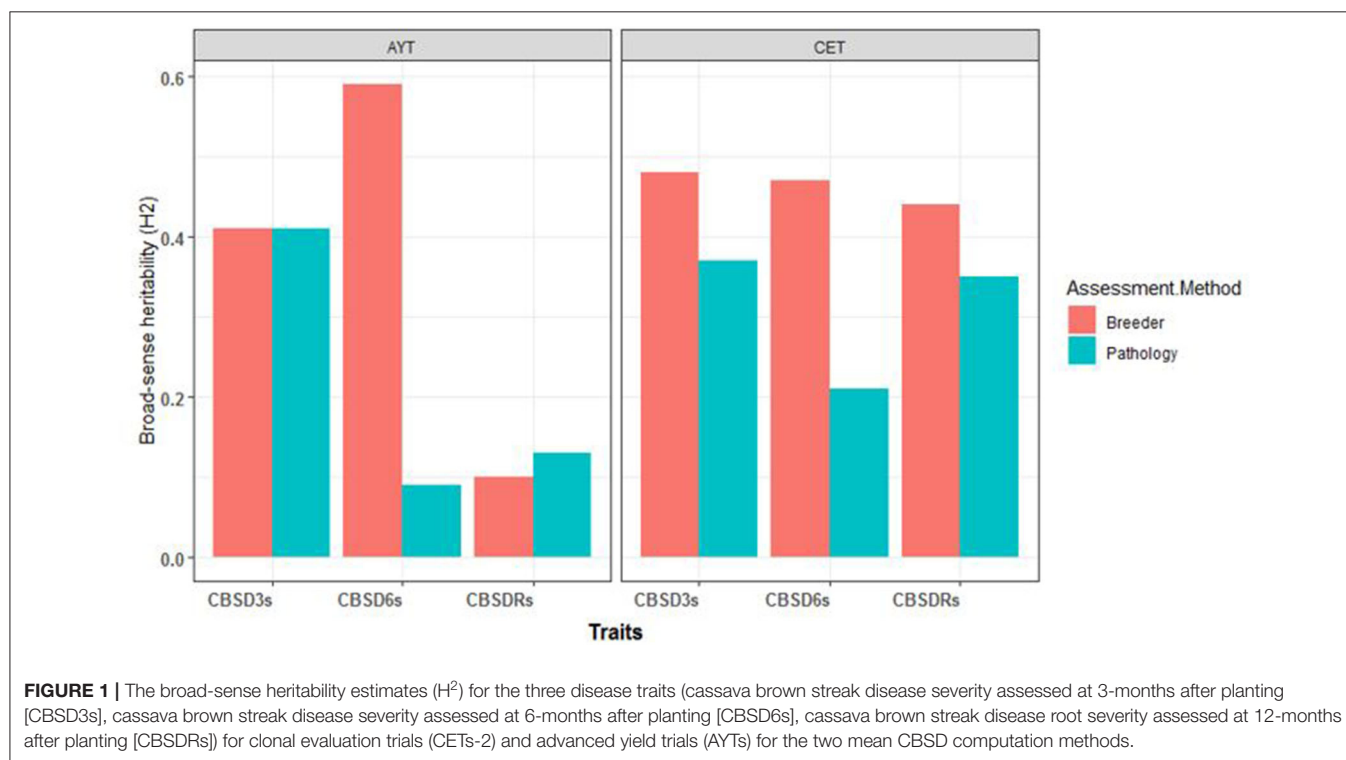
On recognizing the CBSD epidemic in Uganda in the early 2000s, concerted research efforts were initiated to understand the diversity of viruses causing CBSD (Mbanzibwa et al., 2011; Alicai et al., 2016; Ateka et al., 2017; Mbewe et al., 2017), their transmission by the whitefly vector, *Bemisia tabaci* (Maruthi et al., 2005; Omongo et al., 2012; Mugerwa et al., 2018; Ally et al., 2019), and sourcing for resistant genetic materials for breeding (Kanju et al., 2007; Kawuki et al., 2016). More recently, transgenic approaches have also been explored to combat CBSD, but with no officially released genetically transformed plant under cultivation in Uganda to date (Patil et al., 2011; Yadav et al., 2011; Wagaba et al., 2017). Collectively, these research interventions have contributed to our increased understanding and management of CBSD.

A discrepancy remains in the methodologies of CBSD resistance screening, which continues to be refined (Kawuki et al., 2019). In general, in screening for CBSD resistance, plant pathologists assess clone performance based on average foliar infected plants and/or roots, i.e., exclude scores of 1

TABLE 1 | Broad and narrow-sense heritability estimates associated with breeder's and pathologist's CBSD root severity assessment methods.

| Trial location | Seasons | C ₀ clones | Broad-sense heritability | | Narrow-sense heritability | |
|-------------------|---------|-----------------------|--------------------------|---------------|---------------------------|---------------|
| | | | Breeder's | Pathologist's | Breeder's | Pathologist's |
| Mityana | 2015A | 115 | 0.39 | 0.51 | 0.10 | 0.04 |
| Mityana | 2015B | 105 | 0.62 | 0.64 | 0.49 | 0.22 |
| Arua | 2015A | 149 | 0.15 | 0.00 | 0.00 | 0.00 |
| Arua | 2015B | 111 | 0.64 | 0.47 | 0.39 | 0.34 |
| Kasese | 2015A | 116 | 0.26 | 0.06 | 0.21 | 0.06 |
| Kasese | 2015B | 138 | 0.54 | 0.57 | 0.51 | 0.34 |
| Kigumba | 2015A | 147 | 0.49 | 0.61 | 0.25 | 0.09 |
| Kigumba | 2015B | 116 | 0.56 | 0.54 | 0.13 | 0.05 |
| Namulonge | 2015A | 150 | 0.87 | 0.71 | 0.72 | 0.31 |
| Namulonge | 2015B | 113 | 0.79 | 0.71 | 0.55 | 0.45 |
| Serere | 2015A | 123 | 0.68 | 0.64 | 0.64 | 0.63 |
| Serere | 2015B | 112 | 0.71 | 0.58 | 0.70 | 0.56 |
| Lira | 2015A | 149 | 0.47 | 0.44 | 0.25 | 0.22 |
| Lira | 2015B | 108 | 0.67 | 0.48 | 0.54 | 0.27 |
| Mean Heritability | | | 0.56 | 0.49 | 0.39 | 0.25 |

2015A and 2015B, refers to the first (April-May) and second (Aug-Sept) planting seasons.



(Ogwok et al., 2012; Odipio et al., 2014). On the other hand, breeders assess clone performance based on average foliar infected plants and/or roots without excluding the severity scores of 1, i.e., no data are excluded (Kanju et al., 2007; Okul et al., 2018; Kawuki et al., 2019; Ozimati et al., 2019). Certainly,

the methods have varying sampling sizes, hence introducing sampling errors or biases. This study aimed at comparing the two CBSD severity assessment methods based on the heritability estimates and the relative ranking of clones at different trial stages.

TABLE 2 | Pearson correlation coefficients of the 50 clones evaluated at CETs-2 and AYT.

| Traits | BLUPs | | Raw phenotype | |
|--------|--------------------|--------------------|--------------------|--------------------|
| | Breeder's | Pathologist's | Breeder's | Pathologist's |
| CBSD3s | 0.40** | 0.29** | 0.37** | 0.19 ^{ns} |
| CBSD6s | 0.37** | 0.20 ^{ns} | 0.36** | 0.26* |
| CBSDRs | 0.02 ^{ns} | 0.03 ^{ns} | 0.05 ^{ns} | 0.11 ^{ns} |
| S.I | 0.20 ^{ns} | 0.06 ^{ns} | 0.21 ^{ns} | 0.09 ^{ns} |

*, **Significant correlation at $p \leq 0.05$ and 0.01 , respectively.

^{ns}, non-significant correlations coefficients; CBSD3s, cassava brown streak disease severity scored 3 months after planting; CBSD6s, cassava brown streak disease severity scored 6 months after planting; CBSDRs, cassava brown streak disease root severity scored at 12-months harvest; S.I, selection index values for the three cassava brown streak traits; BLUPs, best linear unbiased predictors for clones.

TABLE 3 | Ranking of top 10 C₁ clones by their indexed BLUPs values for the two CBSDs.

| Breeder's | | Pathologist's | |
|---------------|--------------|---------------|--------------|
| CETs-2 | AYTs | CETs-2 | AYTs |
| UG15F190P001* | UG15F190P001 | UG15F262P513 | UG15F265P001 |
| UG15F262P513 | UG15F079P011 | UG15F190P001* | UG15F312P003 |
| UG15F170P507* | UG15F140P003 | UG15F017P003 | UG15F190P001 |
| UG15F176P004 | UG15F196P004 | UG15F177P016 | UG15F249P007 |
| UG15F201P517 | UG15F176P502 | UG15F170P507 | UG15F047P010 |
| UG15F079P011* | UG15F177P016 | UG15F306P028 | UG15F044P009 |
| UG15F176P502* | UG15F044P009 | UG15F176P004 | UG15F169P507 |
| UG15F017P003 | UG15F170P507 | UG15F222P038 | UG15F158P005 |
| UG15F209P001 | UG15F222P038 | UG15F361P510 | UG15F140P001 |
| UG15F302P513 | UG15F312P003 | UG15F154P005 | UG15F196P004 |

CETs-2, clonal evaluation trial (C₁); AYT, advanced yield trials.

*Overlapping clones at CET and AYT.

Heritability Estimates of CBSD Foliar and Root Necrosis for the Two Assessment Methods

According to Bernardo (2003), the broad- and narrow-sense heritability estimates are critical for selection decisions. The comparison of heritability estimates across CETs-1 revealed higher heritability estimates for the method of breeders for CBSD root severity assessment than that for the method of pathologists, with the highest plot-based broad-sense ($H^2 = 0.87$) and narrow-sense ($h^2 = 0.72$) heritability estimates recorded for Namulonge trial in 2015A. In a recent study by Kawuki et al. (2019), a minimum number of 30 roots per plot were recommended to obtain the meaningful assessment of CBSD root necrosis. A notable difference between the CBSD assessment methods of breeders and pathologists is that the former uses sample sizes larger (i.e., includes all roots to obtain plot mean) than the latter (i.e., excludes roots with a severity score of 1). Averaging all root scores per plot possibly explains the higher precision and heritability estimates observed for the CBSD assessment of breeders compared with that for the approach of pathologists with the exclusion of roots scores of 1 (i.e.,

no necrosis). In the same study by Kawuki et al. (2019), the lowest standard error from five CBSD root necrosis assessment methods were associated with trials at Namulonge, supporting early studies qualifying Namulonge as a hot spot for CBSD screening (Kaweesi et al., 2014; Okul et al., 2018). It is not surprising that Namulonge presented the highest heritability estimates in this study, supporting it as a hot spot for CBSD screening. Efforts are currently in place to improve the CBSD phenotyping at the hot spot in Namulonge by the use of imaging technology, which is considered a robust and less subjective screening method. As stated by Bernardo (2003), heritability is an important function in the genetic study of metric character, because it reflects the predictive accuracy and reliability of the phenotypic values. Thus, the highest heritability estimates (i.e., broad sense and narrow sense) for both foliar and CBSD root necrosis recorded from the computation of breeders support the use of this method for efficient selection of CBSD-resistant clones.

Comparing Pearson's Correlation Coefficients for BLUP Estimates of Clone in CETs-2 and AYT for the Two CBSD Assessments Methods

The best linear unbiased predictor (BLUP) pioneered by C.R. Henderson (Piepho et al., 2008) as a procedure for genetic estimation was first used for practical dairy breeding. The BLUP procedure allows for a more accurate estimation of genetic merit of traits in the unbalanced data while accounting for the differences in the amount of data available for each genotype (Bernardo, 2003). In general, the correlation coefficients of BLUP values for CBSD traits of clones that were filtered from CETs-2 (C₁) to AYT (C₁) were low to moderate ($r = 0.02$ – 0.40). However, these correlation coefficients were higher and significant ($p \leq 0.01$) for the mean foliar CBSD computation of breeders than the method of pathologists for clones that made it from CETs-2 to AYT. Ozimati et al. (2019) previously reported a high genetic correlation of 0.70 for root necrosis between measurements at seedling vs. at clonal evaluations. In this study, the low correlation observed between BLUPs values at CETs-2 and AYT for root necrosis could be due to degeneration. Recycling the clones for more than three planting seasons has been reported to cause resistance degeneration due to the buildup of the virus population (Shirima et al., 2017). In fact, to date, no clones have been reported to be immune in the conventional breeding pipeline, except for the recent sources of immunity reported from Latin American germplasm (Sheat et al., 2019). One approach of selecting and advancing clones in face of degeneration due to the virus buildup would be to complement the symptom-based screening with the measurements of virus titer, especially when advancing clones from the mid-to-late stages of selection, i.e., from CET stage onward. However, the high cost per assay is a major limitation to the use of quantitative PCR (q-PCR) for virus screening of a large number of clones, as at CET (i.e., over 600 genotypes; Ogowok et al., 2012; Kaweesi et al., 2014; Okul et al., 2018). Through international collaboration with Plant Virus Department, Leibniz Institute DSMZ-German Collection of Microorganism and Cell Culture, Braunschweig,

Germany, a cheap and rapid assay is being developed to enable the screening of large entries. Nonetheless, the higher correlation coefficients observed between the BLUP values of clones in CETs-2 and AYT for mean CBSD computation of breeders than for the approach of pathologists support the use of the assessment methods of breeders for a more effective selection of resistant clones.

Ranking of Clones by Their Indexed BLUPs for the Two CBSD Averaging Methods

In a recent study by Kawuki et al. (2019), to evaluate the alternative methods for assessing CBSD root necrosis, 256 clones were ranked using their BLUPs for five CBSD assessment methods. The comparison of the top 15 resistant clones ranked across the CBSD assessment methods showed one overlapping clone for all the five CBSD root necrosis assessment methods (Kawuki et al., 2019). In this study, ranking of the top 10 resistant clones from CETs-2 and AYT revealed four clones featuring at both evaluation stages for the CBSD assessment method of breeders compared with only a single clone that overlapped for the approach of pathologists. Four clones overlapping at CETs-2 and AYT for breeders mean CBSD computation relative to a single clone for pathologists assessment method, further supports the use of breeders-derived phenotypes to guide selection decisions.

CONCLUSION

This study provides insights into CBSD necrosis assessment as performed by the methods of breeders and pathologists that remarkably differ in how the mean severities are computed. Based on the heritability estimates and the number of clones that were filtered, it was evident that computing mean CBSD for the entire number of roots from a plot was more reliable

compared with cases where roots with severity scores of 1 were excluded.

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

AO conceived the original manuscript idea, analyzed the data, and wrote the manuscript. RK, TA, WE, J-LJ, and CE reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Next Generation Cassava Breeding Project (INV-007637) through funding from Bill and Melinda Gates Foundation and by the International Department of the United Kingdom.

ACKNOWLEDGMENTS

We thank NaCRRI field and laboratory technical staff for their support during data collection and sampling of leaves for DNA extraction, respectively.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.648436/full#supplementary-material>

REFERENCES

- Alicai, T., Ndunguru, J., Sseruwagi, P., Tairo, F., Okao-Okuja, G., Nanvubya, R., et al. (2016). Cassava brown streak virus has a rapidly evolving genome: implications for virus speciation, variability, diagnosis and host resistance. *Sci. Rep.* 6:36164. doi: 10.1038/srep36164
- Alicai, T., Omongo, C. A., Maruthi, M. N., Hillocks, R. J., Baguma, Y., Kawuki, R., et al. (2007). Re-emergence of cassava brown streak disease in Uganda. *Plant Dis.* 91, 24–29. doi: 10.1094/PD-91-0024
- Ally, H. M., El Hamss, H., Simiand, C., Maruthi, M. N., Colvin, J., Omongo, C. A., et al. (2019). What has changed in the outbreaking populations of the severe crop pest whitefly species in cassava in two decades. *Sci. Rep.* 9, 1–13. doi: 10.1038/s41598-019-50259-0
- Ateka, E., Alicai, T., Ndunguru, J., Tairo, F., Sseruwagi, P., Kiarie, S., et al. (2017). Unusual occurrence of a DAG motif in the Ipomovirus Cassava brown streak virus and implications for its vector transmission. *PLoS ONE* 12:187883. doi: 10.1371/journal.pone.0187883
- Bernardo, R. (2003). *Breeding for Quantitative Traits in Plants*. (Woodbury, MN: Stemma Press).
- 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, 1–10. doi: 10.1371/journal.pone.0019379
- FAOSTAT (2019). *Food and Agriculture organization of the United Nations, 2010*. Available online at: <http://www.fao.org/faostat/en/#home> (accessed October 15, 2020).
- Hamblin, M. T., and Rabbi, I. Y. (2014). The effects of restriction-enzyme choice on properties of genotyping-by-sequencing libraries: a study in Cassava (*Manihot esculenta*). *Crop Sci.* 54, 2603–2608. doi: 10.2135/cropsci2014.02.0160
- Hickey, L. T., Hafeez, A. N., Robinson, H., Jackson, S. A., Leal-bertioli, S. C. M., Tester, M., et al. (2017). Breeding crops to feed 10 billion. *Nat. Biotechnol.* 37, 744–754. doi: 10.1038/s41587-019-0152-9
- Hillocks, R. (2004). *Research Protocols for Cassava Brown Streak Virus*. Fao.Org, 1–7.
- Hillocks, R., Maruthi, M., Kulembeka, H., Jeremiah, S., Alacho, F., Masinde, E., et al. (2016). Disparity between leaf and root symptoms and crop losses associated with cassava brown streak disease in four countries in Eastern Africa. *J. Phytopathol.* 164, 86–93. doi: 10.1111/jph.12430
- Hillocks, R. J., and Maruthi, M. N. (2015). Post-harvest impact of cassava brown streak disease in four countries in eastern Africa. *Food Chain* 5, 116–122. doi: 10.3362/2046-1887.2015.008
- Hillocks, R. J., and Thresh, J. M. (2000). Cassava mosaic and cassava brown streak virus diseases in Africa. *Root* 7, 1–8.
- Kanju, E. E., Masumba, E., Masawe, M., Tollano, S., Mulli, B., Mahungu, N., et al. (2007). Breeding cassava for brown streak resistance: regional cassava variety development strategy based on farmer and consumer preferences. *Proc. 13th ISTRC Symp.* 2007, 95–101.
- Kaweesi, T., Kawuki, K., Kyaligonza, V., Baguma, Y., Tusiime, G., and Ferguson, M. E., et al. (2014). Field evaluation of selected cassava genotypes for cassava brown streak disease based on symptom expression and virus load. Field evaluation of selected cassava genotypes for cassava brown streak

- disease based on symptom expression and virus load. *Virol. J.* 2, 11–216. doi: 10.1186/s12985-014-0216-x
- Kawuki, R. S., Esuma, W., Ozimati, A., Kayondo, I. S., Nandudu, L., and Wolfe, M., et al. (2019). Alternative approaches for assessing cassava brown streak root necrosis to guide resistance breeding and selection. 10, 1–13. doi: 10.3389/fpls.2019.01461
- Kawuki, R. S., Kaweesi, T., Esuma, W., Pariyo, A., Kayondo, I. S., Ozimati, A., et al. (2016). Eleven years of breeding efforts to combat cassava brown streak disease. *Breed. Sci.* 66, 560–571. doi: 10.1270/jsbbs.16005
- Legg, J. P., Jeremiah, S. C., Obiero, H. M., Maruthi, M. N., Ndyetabula, I., Okao-Okuja, G., et al. (2011). Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in Africa. *Virus Res.* 159, 161–170. doi: 10.1016/j.virusres.2011.04.018
- Malik, A. I., Kongsil, P., Ou, W., Srean, P., Utsumi, Y., Lu, C., et al. (2020). Cassava breeding and agronomy in Asia: 50 years of history and future directions. *Breed Sci.* 70, 145–166. doi: 10.1270/jsbbs.18180
- Maruthi, M. N., Hillocks, R. J., Mtunda, K., Raya, M. D., Muhanna, M., Kiozia, H., et al. (2005). Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). *J. Phytopathol.* 153, 307–312. doi: 10.1111/j.1439-0434.2005.00974.x
- Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Patil, B. L., Yadav, J. S., Bagewadi, B., et al. (2011). Evolution of cassava brown streak disease-associated viruses. *J. Gen. Virol.* 92, 974–987. doi: 10.1099/vir.0.026922-0
- Mbewe, W., Tairo, F., Sseruwagi, P., Ndunguru, J., Duffy, S., Mukasa, S., et al. (2017). Variability in P1 gene redefines phylogenetic relationships among cassava brown streak viruses. *Virol. J.* 14, 1–7. doi: 10.1186/s12985-017-0790-9
- Mugerwa, H., Seal, S., Wang, H. L., Patel, M. V., Kabaalu, R., Omongo, C. A., et al. (2018). African ancestry of New World, *Bemisia tabaci*-whitefly species. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-20956-3
- Mware, B., Ateka, E., Songa, J., Narla, R. D., Olubaya, F., and Amata, R. (2009). Transmission and distribution of cassava brown streak virus disease in cassava growing areas of Kenya. *J. Appl.* 16:864–870.
- Odipio, J., Ogwok, E., Taylor, N. J., Halsey, M., Bua, A., Fauquet, C. M., et al. (2014). RNAi-derived field resistance to Cassava brown streak disease persists across the vegetative cropping cycle. *GM Crops Food* 5, 16–19. doi: 10.4161/gmcr.26408
- Ogwok, E., Odipio, J., Halsey, J., Gaitán-Solís, E., Bua, A., Taylor, N. J., et al. (2012). Transgenic RNA interference (RNAi)-derived field resistance to cassava brown streak disease. *Mol. Plant Pathol.* 13, 1019–1031. doi: 10.1111/j.1364-3703.2012.00812.x
- Okul, A., Ochwo-Ssemakula, M., Kaweesi, T., Ozimati, A., Mrema, E., Mwale, E. S., et al. (2018). Plot based heritability estimates and categorization of cassava genotype response to cassava brown streak disease. *Crop Prot.* 108, 39–46. doi: 10.1016/j.cropro.2018.02.008
- Omongo, C. A., Kawuki, R., Bellotti, A. C., Alicai, T., Baguma, Y., Maruthi, M. N., et al. (2012). African cassava whitefly, *Bemisia tabaci*, resistance in African and South American Cassava genotypes. *J. Integr. Agric.* 11, 327–336. doi: 10.1016/S2095-3119(12)60017-3
- Ozimati, A., Kawuki, R., Esuma, W., Kayondo, I. S., Wolfe, W., Lozano, R., et al. (2018). Training population optimization for prediction of cassava brown streak disease resistance in West African Clones. *Genes Genomes Genet.* 8, 3903–3913. doi: 10.1534/g3.118.200710
- Ozimati, A., Kawuki, R., Esuma, W., Kayondo, S. I., Pariyo, A., Wolfe, M., et al. (2019). Genetic variation and trait correlations in an East African Cassava breeding population for genomic selection. *Crop Sci.* 59, 460–473. doi: 10.2135/cropsci2018.01.0060
- Patil, B. L., and Fauquet, C. M. (2009). Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol. Plant Pathol.* 10, 685–701. doi: 10.1111/j.1364-3703.2009.00559.x
- Patil, B. L., Legg, J. P., Kanju, E., and Fauquet, C. M. (2015). Cassava brown streak disease: a threat to food security in Africa. *J. Gen. Virol.* 96, 956–968. doi: 10.1099/jgv.0.000014
- Patil, B. L., Ogwok, E., Wagaba, H., Mohammed, I. U., Yadav, J. S., Bagewadi, B., et al. (2011). RNAi-mediated resistance to diverse isolates belonging to two virus species involved in Cassava brown streak disease. *Mol. Plant Pathol.* 12, 31–41. doi: 10.1111/j.1364-3703.2010.00650.x
- Piepho, H. P., Möhring, J., Melchinger, A. E., and Büchse, A. (2008). BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161, 209–228. doi: 10.1007/s10681-007-9449-8
- R Development Core Team (2008). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* (ISBN 3-900051-07-0): 900051. Vienna: R Found. Stat. Comput.
- Sheat, S., Fuerholzner, B., Stein, B., and Winter, S. (2019). Resistance against cassava brown streak viruses from Africa in cassava germplasm from South America. *Front. Plant Sci.* 10:567. doi: 10.3389/fpls.2019.00567
- Shirima, R., Maeda, D., Kanju, E., Caesar, G., Tibazarwa, F., and Legg, J. (2017). Absolute quantification of cassava brown streak virus mRNA by real-time qPCR. *J. Virol. Methods* 245, 5–13. doi: 10.1016/j.jviromet.2017.03.003
- Wagaba, H., Beyene, G., Aleu, J., Odipio, J., Okao-Okuja, G., Chauhan, R. D., et al. (2017). Field level RNAi-mediated resistance to cassava brown streak disease across multiple cropping cycles and diverse East African agro-ecological locations. *Front. Plant Sci.* 7:2060. doi: 10.3389/fpls.2016.02060
- Wolfe, M. D., Del Carpio, D. P., Alabi, O., Ezenwaka, L. C., Ikeogu, U. N., Kayondo, I. S., et al. (2017). Prospects for genomic selection in cassava breeding. *Plant Genome* 10, 1–19. doi: 10.3835/plantgenome2017.03.0015
- Wolfe, M. D., Rabbi, I. Y., Egesi, C., Hamblin, M., Kawuki, R., Kulakow, P., et al. (2016). Genome-wide association and prediction reveals genetic architecture of cassava mosaic disease resistance and prospects for rapid genetic improvement. *Plant Genome* 9, 342–356. doi: 10.3835/plantgenome2015.11.0118
- Yadav, J. S., Ogwok, E., Wagaba, H., Patil, B. L., Bagewadi, B., Alicai, T., et al. (2011). RNAi-mediated resistance to Cassava brown streak Uganda virus in transgenic cassava. *Mol. Plant Pathol.* 12, 677–687. doi: 10.1111/j.1364-3703.2010.00700.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 Ozimati, Esuma, Alicai, Jannink, Egesi and Kawuki. 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.



Pre-emptive Breeding Against Karnal Bunt Infection in Common Wheat: Combining Genomic and Agronomic Information to Identify Suitable Parents

Livinus Emebiri^{1,2*}, Shane Hildebrand¹, Mui-Keng Tan³, Philomin Juliana⁴, Pawan K. Singh^{4*}, Guillermo Fuentes-Davila⁵ and Ravi P. Singh⁴

¹ NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia, ² Graham Centre for Agricultural Innovation (NSW Department of Primary Industries and Charles Sturt University), Wagga Wagga, NSW, Australia, ³ NSW Department of Primary Industries, Menangle, NSW, Australia, ⁴ International Maize and Wheat Improvement Center, Mexico City, Mexico, ⁵ Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Cd. Obregón, Mexico

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Devanna B N,
National Rice Research Institute
(ICAR), India
Yong Suk Chung,
Jeju National University, South Korea

*Correspondence:

Pawan K. Singh
Pk.Singh@cgiar.org
Livinus Emebiri
livinus.emebiri@dpi.nsw.gov.au

Specialty section:

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

Received: 04 March 2021

Accepted: 18 May 2021

Published: 29 July 2021

Citation:

Emebiri L, Hildebrand S, Tan M-K, Juliana P, Singh PK, Fuentes-Davila G and Singh RP (2021) Pre-emptive Breeding Against Karnal Bunt Infection in Common Wheat: Combining Genomic and Agronomic Information to Identify Suitable Parents. *Front. Plant Sci.* 12:675859. doi: 10.3389/fpls.2021.675859

Wheat (*Triticum aestivum* L.) is the most widely grown cereal crop in the world and is staple food to half the world's population. The current world population is expected to reach 9.8 billion people by 2050, but food production is not expected to keep pace with demand in developing countries. Significant opportunities exist for traditional grain exporters to produce and export greater amounts of wheat to fill the gap. Karnal bunt, however, is a major threat, due to its use as a non-tariff trade barrier by several wheat-importing countries. The cultivation of resistant varieties remains the most cost-effective approach to manage the disease, but in countries that are free of the disease, genetic improvement is difficult due to quarantine restrictions. Here we report a study on pre-emptive breeding designed to identify linked molecular markers, evaluate the prospects of genomic selection as a tool, and prioritise wheat genotypes suitable for use as parents. In a genome-wide association (GWAS) study, we identified six DARTseq markers significantly linked to Karnal bunt resistance, which explained between 7.6 and 29.5% of the observed phenotypic variation. The accuracy of genomic prediction was estimated to vary between 0.53 and 0.56, depending on whether it is based solely on the identified Quantitative trait loci (QTL) markers or the use of genome-wide markers. As genotypes used as parents would be required to possess good yield and phenology, further research was conducted to assess the agronomic value of Karnal bunt resistant germplasm from the International Maize and Wheat Improvement Center (CIMMYT). We identified an ideal genotype, ZVS13_385, which possessed similar agronomic attributes to the highly successful Australian wheat variety, Mace. It is phenotypically resistant to Karnal bunt infection (<1% infection) and carried all the favourable alleles detected for resistance in this study. The identification of a genotype combining Karnal bunt resistance with adaptive agronomic traits overcomes the concerns of breeders regarding yield penalty in the absence of the disease.

Keywords: Karnal bunt resistance, *Tilletia indica*, wheat, *Triticum aestivum*, genome-wide association study, GWAS, genomic prediction, grain yield

INTRODUCTION

Wheat (*Triticum aestivum* L.) is the most widely grown cereal crop on the planet, a staple of the world economy, supplying one fifth of calories consumed by people each day (Anonymous, 2020). The current world population of about 7.7 billion is expected to increase and reach 9.8 billion people by 2050. To accommodate the increased demand for food, annual cereal production will need to rise by about 60–70% from the current level of 2.8 billion tonnes. For various reasons, however, production is not expected to keep pace with demand in developing countries, and their net imports of cereals are projected to more than double from 135 million metric tonnes in 2008/2009 to 300 million metric tonnes in 2050 (Food Agriculture Organization 2009). This gap can be bridged by increased imports, and significant opportunities exist for traditional grain exporters, including Australia, to produce and export greater amounts of wheat over the next few decades (Linehan et al., 2012). Karnal bunt, a disease caused by the fungus *Tilletia indica* Mitra [syn. *Neovossia indica* (Mitra) Mundkur], is a threat to grain export (Joshi et al., 1983), due to its use as a non-tariff trade barrier by several wheat-importing countries (Beattie and Biggerstaff, 1999). The disease has minimal impact on wheat grain yield (Warham, 1986; Murray and Brennan, 1998) but the infected grains exude an unpleasant, rotten fish odour due to a chemical (trimethylamine) produced by the fungal spores (Mitra, 1935). Trimethylamine is associated with multiple diseases in humans, including renal disorders, cancer, obesity, and cardiovascular diseases (Chhibber-Goel et al., 2016).

Control of this disease is difficult because teliospores of the fungus are resistant to physical and chemical factors (Fuentes-Dávila et al., 2018), the fungus causes local infections (Fuentes-Dávila, 1996), and teliospores may remain dormant for more than 32 months (Babadoost et al., 2004). The cultivation of resistant varieties remains the most cost-effective approach to manage the threat of incursions into countries free of the disease (Singh et al., 2007; Emebiri et al., 2019a). Sources of resistance have been identified in the wild relative of wheat, *Aegilops tauschii* (Chhuneja et al., 2008), and in synthetic hexaploid wheat (Mujeeb-Kazi et al., 2006), but resistance in common wheat is limited (Fuentes-Dávila and Rajaram, 1994), and as such, progress in breeding resistant varieties has remained modest. In most wheat-exporting countries that are free of the disease, there are no breeding efforts due to cost burdens and the low return on investments, which in the absence of an incursion, is zero (White et al., 2016). Availability of molecular markers closely linked to resistance genes could be incentivising, as it has the potential to improve selection (Singh et al., 2012; Emebiri et al., 2019b), but efforts in the past have also been modest. Quantitative trait loci (QTL) associated with Karnal bunt resistance in common wheat have been identified in the past (Nelson et al., 1998; Singh et al., 2003; Singh et al., 2007, 2012; Kumar et al., 2007, 2015; Kaur et al., 2016), but these studies were based on a small number of restriction fragment length polymorphisms (RFLP) and PCR-based simple sequence repeats (SSRs). Recently, the use of high-density single nucleotide polymorphism (SNP) arrays in genome-wide association studies (GWAS) have been reported

(Brar et al., 2018; Emebiri et al., 2019b; Gupta et al., 2019; Singh et al., 2020), which offers new opportunities for marker-assisted selection (MAS). However, the focus of many plant breeders has now shifted from the use of MAS to the application of genomic selection.

Genomic selection, first introduced by Meuwissen et al. (2001), would be an attractive tool for pre-emptive breeding against exotic pathogens, as it would reduce the challenges of phenotyping (Poland and Rutkoski, 2016). Genomic selection is a two-stage process in which whole-genome markers are used to predict genomic estimated breeding value (GEBV) of individuals in a population, and then selection decisions are made on the basis of these GEBVs (Meuwissen et al., 2001). In the best-case scenario, breeders can select the best performing genotypes from the population for use in their crossing block, without the need to phenotype the plants themselves. The potential for genomic selection has yet to be evaluated for Karnal bunt resistance in common wheat. The prediction accuracy depends on the trait's heritability, and for Karnal bunt resistance, the estimates are quite high (ranging from 0.75 to 0.91) (Brar et al., 2018; Emebiri et al., 2019a; Gupta et al., 2019; Singh et al., 2020) due to the well-established protocol for disease screening (Fuentes-Dávila et al., 1995).

The International Maize and Wheat Improvement Center (CIMMYT), Mexico, develops novel common wheat germplasm carrying Karnal bunt resistance genes (Singh et al., 2016). Some of the lines were derived from crosses that include Munal#1 (now released as Super 172) and synthetic hexaploids (Mujeeb-Kazi et al., 2006) as parents, and some were developed from backcrosses to commercial varieties, such as Batavia, and Pastor. The lines are important pre-emptive breeding tools to prevent the spread of this quarantined disease into countries that are currently disease-free. However, many other variables are involved in grower uptake of new varieties, with grain yield as the ultimate determinant of which variety the farmer will grow in any given season. In the absence of a disease pressure, genetic resistance may in fact become a liability (yield penalty), as demonstrated in numerous studies (Brown, 2002; Ning et al., 2017). Sharp et al. (2002), for instance, observed that while the *Wsm1* gene in wheat provided the most effective resistance to wheat streak mosaic virus, a mean yield reduction of 21% occurred in the absence of the virus. The wheat stem rust resistance gene, *Sr26*, has a 9% yield penalty (Brown, 2002), and the barley (*Hordeum vulgare*) *mlo* resistance gene has a 4.2% yield penalty (Jorgensen, 1992). This is because genetic resistance is an on-going process, and plants expend metabolic energy that might otherwise be converted to yield. In the absence of the pathogen, existence of a yield penalty for Karnal bunt resistance will outweigh the value of the resistance gene (Oliver et al., 2014; Ning et al., 2017), and breeders will be further discouraged from adopting and using improved germplasm in their programmes for fear of upsetting the established phenology and yield profiles.

The key to pre-emptive breeding would be to provide breeders with a package of molecular markers and resistance genes in genetic backgrounds that will not upset established yield and phenology profiles, as there is no point selecting less susceptible varieties if there is an opportunity cost of lower yield without

disease. In this paper, we report a research on pre-emptive breeding for Karnal bunt resistance designed to identify linked molecular markers, assess prospects of genomic selection as a tool, and prioritise wheat genotypes suitable for use as parents. To identify such genotypes, we performed field experiments over two years to compare their agronomic values with those of reference, commercial varieties.

MATERIALS AND METHODS

Plant Materials

The germplasm materials consisted of 242 genotypes, made up of 177 bread wheat varieties, 8 durum wheat, 11 triticale, and 46 Karnal bunt-resistant germplasm lines (KBRL). The KBRL were developed at the CIMMYT, and imported into Australia through the CIMMYT-Australia-ICARDA Germplasm Evaluation (CAIGE) suite of projects.¹ The wheat varieties represent parents used in breeding programmes, historical varieties, and current commercial varieties that are still being cultivated. These were mainly bred in Australia, but some originated from the United States, Brazil, Canada, China, Mexico, New Zealand, and India, providing a global resource for genetic analysis. The bread wheat lines include Super172 (synonym Munal-#1), used as the resistant check, and the highly susceptible Indian wheat variety WL-711 (synonym WL-711-0IND) used as a susceptible check. The names of the varieties, year of release and pedigrees are listed in **Supplementary Table 1**.

Disease Phenotyping

Phenotypic data on Karnal bunt resistance collected from Australian wheat varieties and CIMMYT advanced breeding lines were used. The data for the Australian varieties were derived from field experiments (Emebiri et al., 2019a) conducted during three consecutive cropping seasons (2014–2015, 2015–2016, and 2016–2017), at the Norman E. Borlaug Experimental Station, the CIMMYT, Obregon. The data on CIMMYT breeding lines, collected over three planting dates, were kindly provided by Dr. Ravi P. Singh as part of the materials delivered through CAIGE project. In these data, Karnal bunt resistance was calculated as the percentage of infected grains in each ear (Fuentes-Dávila and Rajaram, 1994), but to rate the genotypes consistently across data sets, those with infection levels of 0–2.5% were rated as resistant, 2.6–5% as moderately resistant, 5.1–10% as moderately susceptible and greater than 10% as susceptible (Gaudet et al., 2001).

Genotyping

Genomic DNA was isolated from the leaves of individual lines as described in Tan et al. (2015) and genotyped using DArT-Seq technology (Diversity Arrays Technology Pty Ltd., Australia). The polymorphisms were scored as binary data (0/1), indicating the presence/absence of SNP in the genome of each sample. The DArTseq data were filtered for quality, first by removing duplicates and monomorphic markers; then by retaining markers

on the basis of CallRate (≥ 0.95), reproducibility (≥ 0.95), minor allele frequency (≥ 0.05), and percent missing data ($\leq 15\%$). The final molecular marker data set comprised of 8,012 loci scored on 177 hexaploid genotypes. All heterozygotes were treated as missing data, and the corresponding values were imputed using the Random Forest regression method in R package (Stekhoven and Bühlmann, 2012).

Genetic Structure and Linkage Disequilibrium

Genetic structure was analysed using algorithms implemented in the *adeigenet* package (Jombart, 2008). First, we ran the *snappclust* function to select the optimal number of genetic groups, based on a statistical measure of goodness of fit, the Bayesian Information Criterion (BIC). Then, a discriminant analysis of principal components (DAPC) was applied, which combined PCA with discriminant analysis to maximise between-group differences while minimising the within-group variation (Jombart et al., 2010).

Linkage disequilibrium (LD) (statistical association between allelic variants) was calculated in *plink* v1.9 (Purcell et al., 2007) as the squared correlation coefficient (r^2) between alleles at pairs of loci within each chromosome. The analyses were carried out with a molecular data set that was thinned down evenly across the genome to a window size of 8 kb. The decay of LD over genetic distance was examined by plotting the pair-wise LD against distance, and fitting a decay curve, established by square root transformation of the predicted LD values calculated according to Andreescu et al. (2007). The background r^2 value was calculated as the 95th percentile of all LD values between markers located on different chromosomes, assumed to be unlinked (Brescaglio and Sorrells, 2006).

Genome-Wide Association Analysis

Genome-wide analyses were performed with the R package, *lmem.gwaser* (Gutierrez et al., 2016), according to the Kinship model, which had a lambda value of 1.03. It can be described as follow:

$$y = x\beta + zu + \varepsilon,$$

where y is the observed phenotype, x is the molecular marker score matrix, β is the vector of marker allelic effects, z is an incidence matrix, u is a vector of random polygene background effects with $\text{Var}(u)$ being $2KV_G$ (K = Kinship coefficients and V_G = genetic variance), and ε is a vector of random experimental error.

We adjusted observed P -value for multiple testing using two methods: the method of Li and Ji (2005), which is based on the effective number of independent tests (alpha level of 0.05) and the false discovery rate (FDR) method of Benjamini and Hochberg (1995). The method of Li and Ji (2005) was implemented in the *lmem.gwaser* package but FDR was calculated in the R function, *p.adjust()*. Allelic effects and proportion of phenotypic variance (R^2 -values) explained by significant markers were derived from simple linear regression analyses, with $R^2 = \text{SS}_{\text{reg}}/\text{SS}_{\text{tot}}$, where SS_{reg} is the regression sum of squares and SS_{tot} is the total sum of squares.

¹<http://caigeproject.org.au/>

Physical Mapping

Significant markers were assigned to physical positions in megabase pairs (Mbp) by nucleotide BLAST (BLASTN) search (E-value threshold = $1E-5$) against the IWGSC RefSeq v1.0 Chinese Spring assembly,² using the marker sequence for query. High-confidence candidate genes closely matching the marker sequence were obtained in a window size of estimated LD each side of the marker. The results were further refined with the *JBrowse* tool (Buels et al., 2016) to identify nearby wheat expressed sequence tags (wEST), and this allowed assigning the markers to physical bin positions on the deletion maps of the Chinese Spring cultivar.

Prediction of Karnal Bunt Resistance

Two scenarios were considered for genomic prediction: (1) the use of only the markers identified in GWAS analysis (QGBLUP), analogous to marker-assisted selection strategy and (2) the use of genome-wide markers to predict the performance of individuals for which genotypic data is available, but not the phenotypes. The analysis was carried out using the genomic best linear unbiased prediction (GBLUP) model (Meuwissen et al., 2001), in which the G-matrix was calculated using either the six significant markers identified in GWAS, or the genome-wide markers, depending on the approach. In both cases, accuracies were determined from a fivefold cross-validation scheme, in which 80% of the genotypes were randomly assigned to a training set (TRN) and the remaining 20% to a testing set (TSN). This was repeated 100 times, and for each repeat, the individuals in the TRN and TSN set were randomly re-sampled, the phenotypes of individuals in the TST set were masked, and then predicted based on the TRN set. Genomic prediction accuracy was calculated as the Pearson's correlation between the actual and the predicted phenotypes of the lines in the TSN set.

Agronomic Assessment

Field experiments were carried out in 2015 and 2016 cropping seasons to assess the agronomic value of Karnal bunt-resistant lines. For this study, 37 of the Karnal bunt resistant lines from CIMMYT were used. Seven commercial wheat varieties were included as reference genotypes. These included Super172 (syn. Munal#1), Axe, Mace, Rosella, Scout, Suntop, and Waagan. Axe was released in 2007 and is a very early maturing wheat that is suited for short growing seasons, while Mace, released in 2008, has broad adaptation, with consistently high yield under a wide range of conditions. Rosella is a widely adapted winter wheat used for dual-purpose grazing, while Suntop was released in 2011 as a main season line, with high and stable yields from low to high yield potential areas. Both Scout and Waagan were released in 2009. Scout is a mid-season maturity variety with low screenings and high test weight, and Waagan is a very early maturing spring wheat, with high yield potential in medium/low rainfall environments.

The experiments were conducted at the Wagga Wagga Agricultural Institute, Wagga Wagga NSW, Australia (latitude – 35.05° S, longitude 147.35° E), on a site with well-drained, sandy

clay loam soil with a greyish brown colour. The experiments were arrayed in a row-column, *p*-rep design (Cullis Brian et al., 2006), with experimental units (plots) measuring 7.5 m² in area (six rows with 30 cm spacing, 6 m long, trimmed to 5 m prior to harvest). Plots were sown with a tractor-mounted Seeder, at a rate of 60-g seeds per plot. All experiments were fertilised at the time of sowing with monoammonium phosphate at the rate of 100 kg/ha, and standard operational procedures (irrigation, weed, pest/disease control) were applied.

Statistical Analysis of Agronomic Data

Data on the following agronomic traits were collected: emergence counts (number of plants per plot), flowering date (50% awn emergence), plant height (height from soil to tip of the awns), NDVI (at anthesis using the GreenSeeker) and grain yield (weight of the uncleaned seed weight from machine harvests per plot). At harvest, the uncleaned grains (300 g) were subsampled and used to collect data on grain size (1,000 grain weight) and grain plumpness (grains retained over a 2.5 mm sieve).

A two stage approach was used for data analysis. In the first stage, each trait within an experiment/year was analysed separately to account for design factors and spatial field variation. This was performed using a mixed linear model framework with spatial corrections for field heterogeneity as implemented in the R package, *SpATS* (Rodríguez-Álvarez et al., 2018). The analytical model included data on seedling emergence (count) per plot as a fixed component to adjust for differences in plant density. *SpATS* uses two-dimensional smoothing surfaces with penalised splines to model the spatial trends within the field and obtain estimates of predicted means. In the second stage, adjusted means for the 2 years were jointly modelled to generate variance components, and a genotype × trait matrix, which was analysed according to the genotype plus genotype-by-environment method, as implemented in *GGEbiplotGUI* (Frutos et al., 2014). Graphical displays of the output were aided by the R package, *ggplot2* (Ginestet, 2011).

RESULTS

Phenotypic Variation

Broad-sense heritability of Karnal bunt resistance, calculated as the ratio of genotypic to phenotypic variance components, was 0.83 ± 0.02 , and for narrow-sense heritability, calculated using a marker-based approach (Covarrubias-Pazaran, 2016), the estimate was relatively high at 0.61 ± 0.14 . These estimates indicated a large contribution of genetic factors to Karnal bunt resistance in the wheat accessions. The average percentage infection in the wheat accessions was 17.5%, with a range of 0.4–51.8%. There were 10 resistant lines, that is, genotypes with seed infection levels of 0–2.5%. These included seven KBRL and three cultivated varieties. Sixteen of the accessions were moderately resistant (% KB infection > 2.5–5%), 46 were moderately susceptible (> 5–10%) and 105 were susceptible (%KB infection > 10%).

²<https://urgi.versailles.inra.fr/blast/>

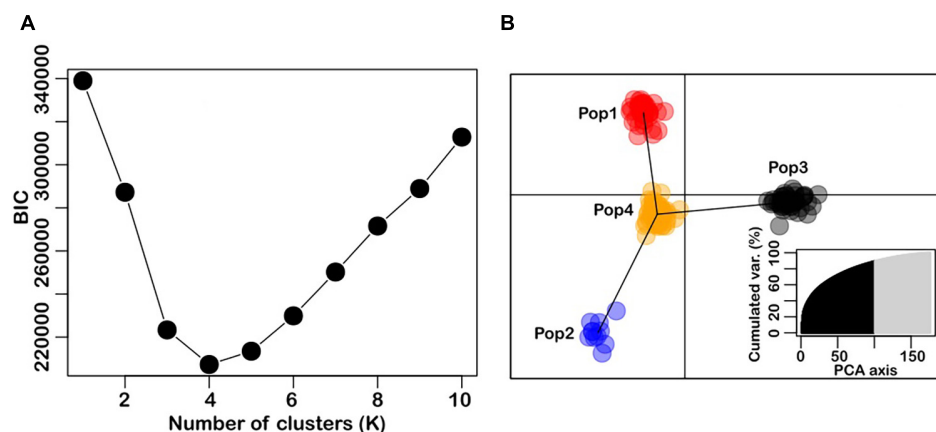


FIGURE 1 | Population structure in the panel of 177 wheat accessions used for the study. Panel **(A)** is the optimal number of clusters identified with the find.cluster function in adegenet (Jombart, 2008). Panel **(B)** is the DAPC results, showing relative positions of individuals and genetic clusters in the discriminant space (inset is the PCA eigenvalues).

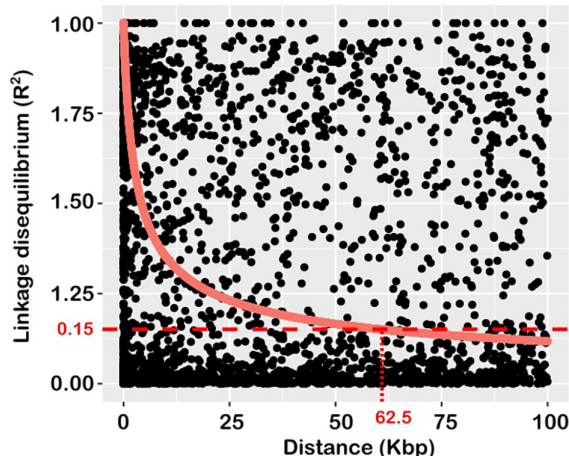


FIGURE 2 | Plot of average linkage disequilibrium (LD) values (r^2) against inter-marker distances over a short (100 kbp) distance to visualise LD decay. The decay curve is the square root transformation of predicted LD values according to Andreescu et al. (2007). The horizontal and vertical dotted lines indicate the baseline r^2 threshold value, and the extent of LD decay, respectively.

Genetic Structure

Genetic structure analysis was performed to determine whether the composition of wheat accessions was structured, that is, differentiated into clusters of closely related individuals, and which individuals belong to which clusters. Graph of BIC values showed a minimum value at $K = 4$, and this was determined to be the optimal number of genetic clusters in the wheat accessions (Figure 1A). The DAPC analysis showed clear separation of the accessions into four genetic clusters (Figure 1B), with sample sizes ranging from 11 to 79. It was noteworthy that the CIMMYT-derived Karnal bunt resistant lines cluster together in Pop3 ($n = 46$), and along with varieties such as Seri-M82, Pastor,

TABLE 1 | Summary of significant markers detected in association mapping of Karnal bunt resistance in common wheat.

| Peak marker | Chr. | Physical position (Mb) | Deletion bin | P-value | FDR-value | Allelic effect | R-squared (%) |
|-------------|------|------------------------|-----------------|----------|-----------|----------------|---------------|
| 2282741 | 1A | 481.52 | 1AL1-0.17-0.61 | 2.08E-05 | 3.33E-02 | 1.10 | 26.91 |
| 1249729 | 2A | 723.62 | C-2AL1-0.85 | 1.45E-04 | 1.16E-01 | 0.92 | 7.58 |
| 1037716 | 3B | 618.02 | 3BL10-0.50-0.63 | 7.64E-07 | 6.12E-03 | -1.04 | 29.47 |
| 993727 | 4A | 719.43 | 4AL4-0.80-1.00 | 4.44E-05 | 5.08E-02 | -1.03 | 26.31 |
| 1128414 | 5A | 618.27 | 5AL17-0.78-0.87 | 6.07E-06 | 2.43E-02 | -1.06 | 27.87 |
| 989877 | 6B | 683.23 | 6BL5-0.40-1.00 | 3.06E-04 | 1.88E-01 | 0.26 | 8.09 |

Nominal P-values were adjusted using the false discovery rate (FDR) method of Benjamini and Hochberg (1995); the explained variation (R-squared) and allelic effects attributable to each marker were derived from simple linear regression analyses.

Genaro-F81, and Veery5, they were separate from the other wheat genotypes. Seri-M82 and Genaro-F81 are semi-dwarf, historical wheat varieties from CIMMYT, and Pastor is derived from a cross involving Ser-M82 as a parent. The genotypes in Pop1 ($n = 41$) and Pop2 ($n = 11$) were Australian-bred wheat varieties, and those in Pop4 ($N = 79$) were a mixture of global wheat genotypes. They include Australian varieties such as Axe, Drysdale and EGA-Burke, the Indian variety WL-711, Canadian varieties such as AC-Domain and its progeny, AC-Snowbird, the Chinese variety, Chuan-Mai-18, the Brazilian variety, Carazinho and the USA variety, Angus.

Linkage disequilibrium (statistical association between allelic variants) and its decay rate were examined using pair-wise combinations of markers genotyped across the 21 wheat chromosomes. The estimate of background LD, calculated from

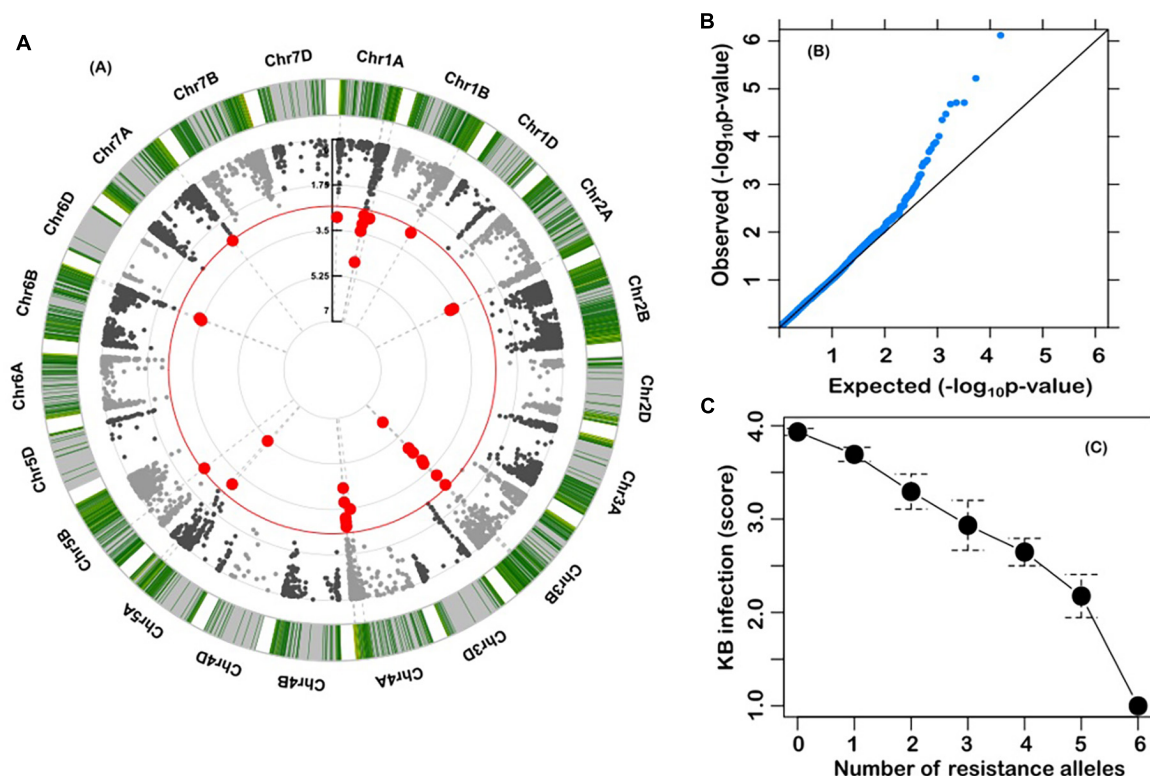


FIGURE 3 | (A) Circular Manhattan plot from genome-wide scan with a mixed linear model. The red line is the significance threshold; **(B)** QQ plots from genome-wide scan. The late separation between observed and expected P-values in the upper left section represents the significant associations; and **(C)** Relationship between number of favourable alleles and Karnal bunt resistance in the wheat accessions.

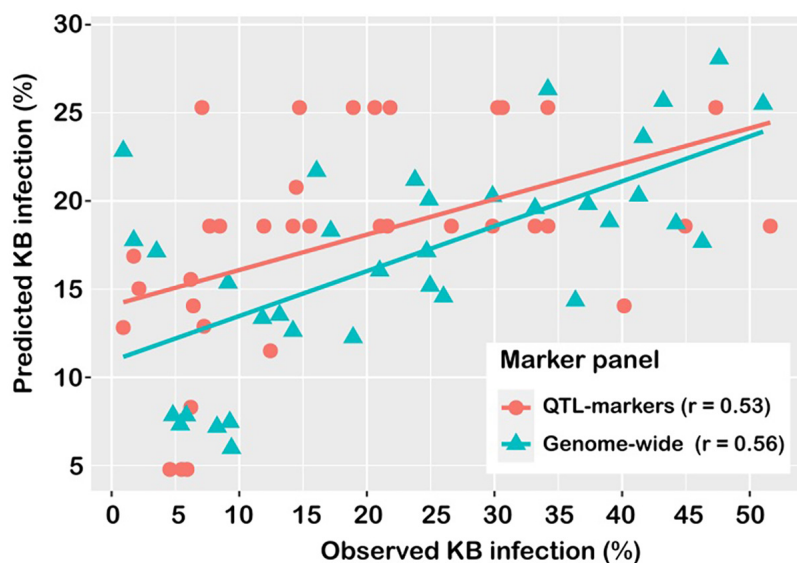


FIGURE 4 | Prediction accuracy of models for Karnal bunt resistance, using markers detected in GWAS analysis (QTL), and genome-wide markers (WG).

r^2 values of unlinked markers was 0.15, which agrees with the value commonly reported for wheat (Joukhadar et al., 2020). This value intersected the LD decay line at 62.5 Kbp (Figure 2), and

this represents the extent of LD in the population used for this study. It represents the mapping resolution of any QTL detected and was used as the confidence interval.

TABLE 2 | Spatially adjusted means of check varieties and CIMMYT-derived, Karnal bunt resistant germplasm.

| Source | Disease rating | Grain yield (t ha ⁻¹) | Flowering date (days) | Plant height (cm) | 1,000 Kernel weight (g) | NDVI | Plump grains (%) |
|-------------------------|----------------|-----------------------------------|-----------------------|-------------------|-------------------------|------|------------------|
| σ^2 genetic | | 1.14 | 22.51 | 68.44 | 42.47 | 0.00 | 17.96 |
| σ^2 residual | | 0.14 | 4.18 | 12.56 | 2.84 | 0.00 | 5.343 |
| Heritability | | 0.94 | 0.92 | 0.92 | 0.97 | 0.87 | 0.87 |
| SE heritability | | 0.08 | 0.11 | 0.11 | 0.04 | 0.16 | 0.16 |
| Adjusted genotype means | | | | | | | |
| Axe | S | 4.75 | 128.64 | 85.72 | 43.90 | 0.57 | 91.83 |
| Mace | S | 5.11 | 128.75 | 91.28 | 41.90 | 0.57 | 89.83 |
| Rosella | S | 5.04 | 135.14 | 97.26 | 38.02 | 0.51 | 85.17 |
| Scout | S | 5.56 | 129.55 | 92.28 | 44.76 | 0.58 | 91.72 |
| Suntop | S | 5.26 | 130.28 | 94.08 | 41.78 | 0.60 | 84.34 |
| Super172 | R | 4.92 | 130.34 | 94.66 | 45.11 | 0.59 | 89.02 |
| Waagan | S | 5.45 | 128.54 | 86.36 | 41.06 | 0.57 | 90.39 |
| ZVS13_312 | MR | 5.35 | 130.31 | 93.75 | 47.87 | 0.59 | 91.02 |
| ZVS13_316 | MS | 4.84 | 129.95 | 91.07 | 47.44 | 0.61 | 91.68 |
| ZVS13_385 | R | 4.69 | 128.34 | 89.51 | 43.93 | 0.57 | 89.94 |
| ZVS13_404 | MS | 5.15 | 129.13 | 97.73 | 44.57 | 0.60 | 90.96 |
| ZVS13_406 | MR | 4.46 | 128.54 | 89.66 | 49.69 | 0.59 | 88.86 |
| ZVS13_441 | MS | 4.79 | 129.29 | 99.56 | 44.67 | 0.61 | 92.24 |
| ZWB10_44 | R | 5.09 | 128.96 | 94.61 | 46.44 | 0.58 | 90.29 |
| ZWB10_76 | MR | 5.36 | 129.40 | 99.08 | 47.03 | 0.60 | 89.76 |
| ZWB11_153 | MS | 4.73 | 129.06 | 99.98 | 49.49 | 0.60 | 91.83 |
| ZWB11_172 | R | 5.07 | 130.97 | 97.13 | 45.16 | 0.57 | 91.37 |
| ZWB11_95 | MR | 5.57 | 130.55 | 94.10 | 44.15 | 0.58 | 90.54 |
| ZWB12_103 | MS | 5.33 | 128.96 | 94.81 | 51.79 | 0.59 | 91.94 |
| ZWB12_121 | MS | 4.75 | 128.68 | 94.32 | 47.88 | 0.60 | 91.19 |
| ZWB12_122 | MS | 5.16 | 129.22 | 95.93 | 47.63 | 0.63 | 89.61 |
| ZWB12_123 | MS | 4.71 | 129.58 | 92.15 | 52.55 | 0.58 | 93.66 |
| ZWB12_124 | MS | 4.93 | 128.26 | 88.84 | 52.70 | 0.57 | 93.69 |
| ZWB12_14 | R | 4.97 | 129.23 | 96.78 | 48.14 | 0.61 | 91.46 |
| ZWB12_147 | MS | 4.65 | 128.43 | 96.26 | 51.82 | 0.61 | 93.46 |
| ZWB12_158 | MS | 4.73 | 128.72 | 101.65 | 47.69 | 0.59 | 92.88 |
| ZWB12_16 | S | 5.14 | 128.40 | 93.55 | 46.60 | 0.58 | 92.34 |
| ZWB12_168 | S | 5.13 | 130.14 | 93.86 | 49.10 | 0.58 | 93.11 |
| ZWB12_18 | MR | 4.57 | 128.84 | 92.68 | 49.05 | 0.58 | 93.02 |
| ZWB12_187 | MR | 5.56 | 130.75 | 99.05 | 48.18 | 0.62 | 91.03 |
| ZWB12_189 | MS | 4.82 | 127.85 | 97.52 | 43.74 | 0.60 | 90.55 |
| ZWB12_194 | MS | 4.79 | 128.53 | 94.78 | 48.26 | 0.59 | 92.55 |
| ZWB12_202 | S | 4.89 | 128.75 | 94.30 | 50.25 | 0.55 | 94.61 |
| ZWB12_219 | MS | 4.75 | 130.45 | 96.43 | 44.21 | 0.52 | 88.04 |
| ZWB12_24 | MS | 4.91 | 129.47 | 98.90 | 47.51 | 0.60 | 92.98 |
| ZWB12_29 | MS | 5.31 | 130.32 | 95.10 | 49.92 | 0.61 | 91.83 |
| ZWB12_30 | R | 5.17 | 129.73 | 100.39 | 50.20 | 0.59 | 90.88 |
| ZWB12_31 | MR | 5.32 | 129.60 | 98.16 | 51.10 | 0.60 | 93.32 |
| ZWB12_4 | MR | 5.20 | 129.48 | 95.07 | 48.22 | 0.59 | 90.91 |
| ZWB12_42 | MS | 5.25 | 128.84 | 91.86 | 48.29 | 0.60 | 92.47 |
| ZWB12_62 | MR | 4.80 | 129.33 | 91.24 | 45.91 | 0.59 | 90.12 |
| ZWB12_63 | MS | 4.73 | 129.11 | 94.99 | 46.19 | 0.58 | 89.93 |
| ZWB12_67 | MS | 5.15 | 129.40 | 94.55 | 46.79 | 0.58 | 90.96 |
| ZWB12_86 | MS | 5.27 | 128.11 | 95.54 | 51.17 | 0.60 | 93.19 |

QTL Identification

There was an evident association between genetic groups and Karnal bunt resistance in the population, as majority of the lines in Pop3 were resistant, and separate from the other groups in the DAPC space (**Figure 1B**). This association of population group with resistance was statistically significant, as determined from a chi-square test of independence ($X^2 = 54.81$, P -value < 0.001), hence, corrective measures were applied to adjust for the potential bias in declaring QTL identification.

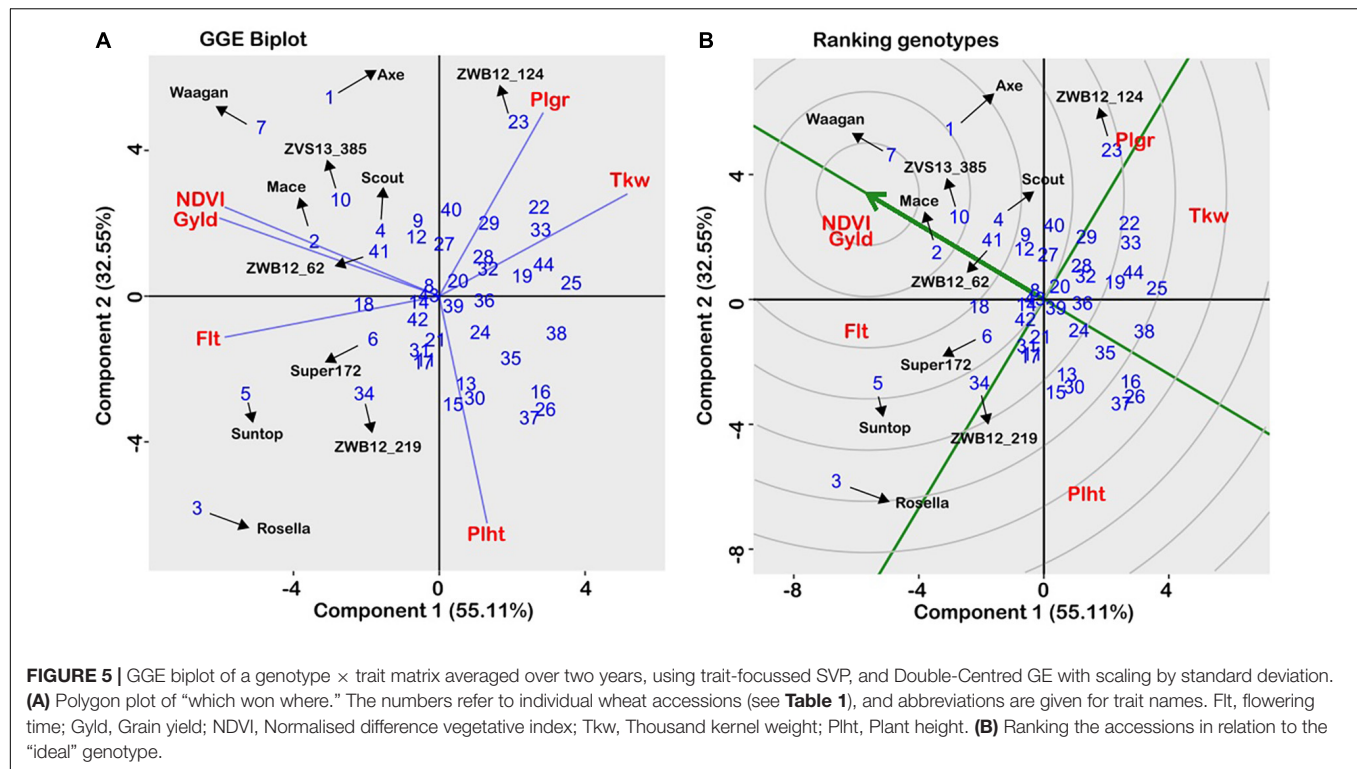
A Kinship-corrected GWAS analysis identified six markers that were significantly associated with Karnal bunt resistance, after controlling for multiple testing using both the genome-wide threshold and FDR criteria (**Table 1** and **Figure 3A**). We compared different mixed models and found the kinship model as the most effective to correct for population structure, as it produced the lowest genomic inflation factor (lambda, $\lambda_{gc} = 1.03$), and the observed P -values showed little deviations from the expected (**Figure 3B**). Surprisingly, all the significant markers were in the A and B genomes, and physically localised to the long arms of chromosomes 1A, 2A, 3B, 4A, 5A, and 6B (**Table 1**). The markers explained a large proportion (7.6–29.5%) of the variation in Karnal bunt resistance, and when favourable alleles were considered, genotypes with a high number of beneficial alleles were completely resistant to Karnal bunt infection (**Figure 3C**).

Genomic Prediction

There was no difference in prediction accuracy between the QGBLUP approach and the whole-genome prediction (GBLUP) approach (**Figure 4**). In the QGBLUP approach, the prediction ability for Karnal bunt resistance averaged 0.53 ± 0.003 , and in the alternate approach of whole genome marker prediction, the accuracy averaged 0.56 ± 0.01 . In effect, genomic prediction using a few, trait-specific markers produced accuracies that compared favourably with those from whole-genome markers.

Agronomic Profiles

In the agronomic experiments, estimates of trait heritability, independent of year and heterogeneous field conditions, were consistently high across traits (**Table 2**), indicating strong genotypic main effects. The adjusted trait means were analysed using the GGE biplot method to allow visual examination of genotype performance across multiple traits, and identification of superior individuals. The biplot captured 87.7% of total variability in the data (**Figure 5**) and is therefore appropriate for visualising the relationships among traits. All traits were equally important, as indicated by the relative length of their vectors. The biplot showed that grain yield was positively related to growth duration and biomass production (acute angles), negatively related to plant height (obtuse angle), and independent of grain size (near right angles). When the “which-won-where” function was used to partition the data into a two-dimensional polygon view, the agronomic traits were grouped into three major sectors (**Figure 5A**): phenology (flowering time/plant height), grain yield (grain yield/NDVI), and grain size (1,000-kernel weight/grain plumpness). Vertex genotypes in each sector



are considered the best/worst for traits within the sector (Yan and Rajcan, 2002). Thus, the late-maturing variety, Rosella, was placed at the apex of the phenology sector, while early maturing varieties, Waagan and Axe, were placed at the vertex of the grain yield sector (**Figure 5A**). These varieties were positioned opposite to the plant height vector, which is consistent with the negative relationship between plant height and grain yield. Of the CIMMYT-derived accessions, ZWB12-124 and ZWB12-147, had the best agronomic values for grain size/plumpness, while ZWB12-158 and ZWB12-30 were the worst for plant height (**Figure 5A**). The mean trait value for all genotypes are presented in **Table 2** to validate the interpretations.

The GGE biplot can also be used to visualise genotype ranking against the “ideal.” The “ideal” is defined as a genotype that combines all favourable attributes, and in **Figure 5B**, the arrow indicates where the ideal genotype should be. Accordingly, the ideal genotype is expected to be high yielding, early maturing and below average in plant height. A performance line passing through the origin is used as a reference, and a genotype closer to the “ideal” is considered more desirable than those further away. As shown in **Figure 5B**, Waagan, followed by Mace and Axe are the more desirable of the check varieties, while Rosella and Suntop were far from the ideal genotype. Of the Karnal bunt-resistant accessions, ZVS13-385 and ZWB12-62 were the closest to the ideal genotype (**Figure 5B**). In particular, the genotype ZVS13-385 was placed within the same concentric ring as Mace, which meant it had similar agronomic attributes. This is relevant information, as Mace is one of the most widely grown varieties in Australia. DNA analysis showed that ZVS13-385 possessed all six of the favourable alleles

identified for Karnal bunt resistance, and therefore would be suitable as a parent for transferring resistance into commercially acceptable backgrounds.

DISCUSSION

In the first part of this study, we sought to dissect the genetic basis of Karnal bunt resistance, as the information is essential for confirming resistance sources, identifying those most suitable as donor parental lines, and designing strategies to accelerate transfer of resistance into commercial cultivars. We identified six DAR-Seq markers, which explained between 7.6 and 29.5% of the observed phenotypic variation and were located at chromosome positions previously reported in the literature (Bishnoi et al., 2020). When BLASTN search was conducted against the IWGSC RefSeq v1.0 Chinese Spring assembly, the most frequently identified putative candidate gene at the QTLs encoded the F-box domain containing proteins. The F-box proteins are a large superfamily that play pivotal roles in host-pathogen interactions through targeting substrates into the degradation machinery (Cao et al., 2008).

The knowledge that Karnal bunt resistance is mediated by multiple genes is supported by previous studies, but this introduces another dimension to the difficulties of breeding for resistance in the absence of the pathogens (Emebiri et al., 2019b). The multi-gene control implies that marker-assisted selection by pyramiding or stacking of favourable alleles may not be successful (Langridge and Waugh, 2019), because interactions among QTL/genes and environmental factors can make substantial

contributions to variation in complex traits such as disease susceptibility (Carlborg and Haley, 2004). As suggested in Emebiri et al. (2019b), new and innovative strategies will be required, and in this study, we assessed the potentials of the method of genomic prediction as a pre-emptive breeding tool. For developing the prediction model, we compared the traditional use of whole-genome markers against the use of a few significant markers identified by GWAS and found the prediction abilities to be comparable (**Figure 4**). This was not surprising, as genomic prediction accuracy is highly dependent on the LD between the genotyped markers and actual causative variants (de Los Campos et al., 2013). The use of significant trait-specific markers was expected to improve genomic prediction, and in fact, prior marker selection has been suggested as a strategy to increase reliability of the genomic estimated breeding values (Brøndum et al., 2015). Rutkoski et al. (2012) reported that in wheat, genomic predictions based on QTL targeted markers for fusarium head blight resistance (deoxynivalenol) alone were higher than predictions based on genome-wide markers. Other researchers have also found higher prediction abilities of the MAS approach over whole-genome prediction (Slavov et al., 2014; Zhao et al., 2014; Boeven et al., 2016), but Gaikpa et al. (2020) found the opposite to be the case. Similarly, while some researchers have found that use of trait-specific markers as fixed factors increased accuracy of genomic prediction (e.g., Daetwyler et al., 2014), others have observed no difference (e.g., Rice and Lipka, 2019). Invariably, this will vary with trait, and the performance of such a prediction model should be explored on a trait-by-trait basis prior to its implementation in a breeding programme (Rice and Lipka, 2019). Karnal bunt resistance in this population showed high heritability (0.83 ± 0.02), hence marker-based prediction accuracies were almost comparable to genome-wide prediction accuracies. This may not be the case in different populations, but the possibility of using a few significant markers for genomic prediction would augur well for pre-emptive breeding against Karnal bunt infection in countries that are free of the disease, where phenotyping would be difficult and the costs for high-density genotyping can be limiting. This is a subject that requires further investigation, as large-scale studies are showing that, in a high LD crop like wheat, high-density genomic coverage has minimal impact on the genomic predictabilities (Juliana et al., 2019).

The identification of parental lines combining Karnal bunt resistance with adaptive agronomic traits is key to pre-emptive breeding, as it addresses breeder's concerns regarding yield penalty in the absence of the disease. Plant breeders use the GGE biplot technique for prioritising genotypes for use as parents in varietal improvement as the regular stability analysis does not provide information on the relative ranking of entries with reference to an ideal genotype (Yan and Kang, 2003). The current research carried out a comprehensive examination of Karnal bunt resistant germplasm from CIMMYT and has identified an ideal genotype, ZVS13_385 (TAM200/PASTOR//TOBA97/3/HEILO), which showed agronomic similarity to the highly successful Australian wheat variety, Mace (Moffat et al., 2015; **Table 2**; **Figure 5B**). Furthermore, ZVS13_385 is phenotypically resistant to Karnal bunt infection (<1% infection), and possessed all

favourable alleles detected for major and minor QTL linked to resistance. This means that it could be used directly as a cultivated variety, or as an ideal genotype for use in the crossing block. We conclude that the identification of a genotype combining Karnal bunt resistance with adaptive agronomic traits negates the concerns of breeders regarding yield penalty in the absence of the disease. Using mathematical modelling, Vyska et al. (2016) showed that even when disease outbreak is uncertain, growing resistant varieties is an optimal strategy for crop protection as it reduces the probability of an outbreak occurring. We add that wide availability of Karnal bunt resistant lines may encourage countries to relax the zero-tolerance regulation that currently exists for Karnal bunt, which is quite costly to implement (Babadoost, 2000; Vocke et al., 2010).

DATA AVAILABILITY STATEMENT

The raw phenotypic data supporting the findings of this article are available and accessible through the CIMMYT-Australia-ICARDA Germplasm Evaluation (CAIGE) suite of related projects (Biosecurity risk – CAIGE (caigeproject.org.au). The genotypic data can be made available by the senior author with no reservations.

AUTHOR CONTRIBUTIONS

LE conceived the study and drafted the manuscript. SH conducted the agronomic experiments. M-KT conducted DNA extractions and supervised the genotyping work. PJ assisted with data analyses. PS, GF-D, and RS conducted the Karnal bunt screening experiments. All authors read and revised the manuscript.

FUNDING

This study was funded jointly by grants from the Grains Research & Development Corporation (GRDC), The NSW Department of Primary Industry (NSW DPI, Australia), and the International Maize and Wheat Improvement Center (CIMMYT, Mexico) under Grant Number DAN00174.

ACKNOWLEDGMENTS

We thank the CGIAR Research Program on Wheat for their support. We also thank all field technicians involved in generating the field and laboratory data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.675859/full#supplementary-material>

REFERENCES

- Andreescu, C., Avendano, S., Brown, S. R., Hassen, A., Lamont, S. J., and Dekkers, J. C. (2007). Linkage disequilibrium in related breeding lines of chickens. *Genetics* 177, 2161–2169. doi: 10.1534/genetics.107.082206
- Anonymous (2020). *The Old, the New and the Newt*. Norwich: John Innes Centre, Norwich Research Park.
- Babadoost, M. (2000). Comments on the zero-tolerance quarantine of Karnal bunt of wheat. *Plant Dis.* 84, 711–712. doi: 10.1094/pdis.2000.84.7.711
- Babadoost, M., Mathre, D. E., Johnston, R. H., and Bonde, M. R. (2004). Survival of teliospores of *Tilletia indica* in soil. *Plant Dis.* 88, 56–62. doi: 10.1094/pdis.2004.88.1.56
- Beattie, B. R., and Biggerstaff, D. R. (1999). Karnal bunt: a wimp of a disease. but an irresistible political opportunity. *Choices* 14:5.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bishnoi, S. K., He, X., Phuke, R. M., Kashyap, P. L., Alakonya, A., Chhokkar, V., et al. (2020). Karnal bunt: a re-emerging old foe of wheat. *Front. Plant Sci.* 11:569057. doi: 10.3389/fpls.2020.569057
- Boeven, P. H., Longin, C. F., Leiser, W. L., Kollers, S., Ebmeyer, E., and Würschum, T. (2016). Genetic architecture of male floral traits required for hybrid wheat breeding. *Theor. Appl. Genet.* 129, 2343–2357. doi: 10.1007/s00122-016-2771-6
- Brar, G. S., Fuentes-Dávila, G., He, X., Sansaloni, C. P., Singh, R. P., and Singh, P. K. (2018). Genetic mapping of resistance in hexaploid wheat for a quarantine disease: Karnal bunt. *Front. Plant Sci.* 9:1497. doi: 10.3389/fpls.2018.01497
- Breseghello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177. doi: 10.1534/genetics.105.044586
- Brøndum, R. F., Su, G., Janss, L., Sahana, G., Gulbrandtsen, B., Boichard, D., et al. (2015). Quantitative trait loci markers derived from whole genome sequence data increases the reliability of genomic prediction. *J. Dairy Sci.* 98, 4107–4116. doi: 10.3168/jds.2014-9005
- Brown, J. K. (2002). Yield penalties of disease resistance in crops. *Curr. Opin. Plant Biol.* 5, 339–344. doi: 10.1016/s1369-5266(02)00270-4
- Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., et al. (2016). JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* 17:66.
- Cao, Y., Yang, Y., Zhang, H., Li, D., Zheng, Z., and Song, F. (2008). Overexpression of a rice defense-related F-box protein gene OsDRF1 in tobacco improves disease resistance through potentiation of defense gene expression. *Physiol. Plant.* 134, 440–452. doi: 10.1111/j.1399-3054.2008.01149.x
- Carlborg, Ö., and Haley, C. S. (2004). Epistasis: too often neglected in complex trait studies? *Nat. Rev. Genet.* 5, 618–625. doi: 10.1038/nrg1407
- Chhibber-Goel, J., Gaur, A., Singhal, V., Parakh, N., Bhargava, B., and Sharma, A. (2016). The complex metabolism of trimethylamine in humans: endogenous and exogenous sources. *Expert Rev. Mol. Med.* 18:e8.
- Chhuneja, P., Kaur, S., Singh, K., and Dhaliwal, H. S. (2008). Evaluation of *Aegilops tauschii* (Coss.) germplasm for Karnal bunt resistance in a screen house with simulated environmental conditions. *Plant Genet. Resour.* 6, 79–84. doi: 10.1017/s1479262108982654
- Covarrubias-Pazarán, G. (2016). Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS One* 11:e0156744. doi: 10.1371/journal.pone.0156744
- Cullis Brian, R., Smith, A. B., and Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11, 381–393. doi: 10.1198/108571106x154443
- Daetwyler, H. D., Bansal, U. K., Bariana, H. S., Hayden, M. J., and Hayes, B. J. (2014). Genomic prediction for rust resistance in diverse wheat landraces. *Theor. Appl. Genet.* 127, 1795–1803.
- de Los Campos, G., Vazquez, A. I., Fernando, R., Klimentidis, Y. C., and Sorensen, D. (2013). Prediction of complex human traits using the genomic best linear unbiased predictor. *PLoS Genet.* 9:e1003608. doi: 10.1371/journal.pgen.1003608
- Emebiri, L., Singh, P. K., Tan, M. K., Fuentes-Dávila, G., He, X., and Singh, R. P. (2019a). Reaction of Australian durum, common wheat, and triticale genotypes to Karnal bunt (*Tilletia indica*) infection under artificial inoculation in the field. *Crop Pasture Sci.* 70, 107–112. doi: 10.1071/cp18235
- Emebiri, L., Singh, S., Tan, M. K., Singh, P. K., Fuentes-Dávila, G., and Ogonnaya, F. (2019b). Unravelling the complex genetics of Karnal bunt (*Tilletia indica*) resistance in common wheat (*Triticum aestivum*) by genetic linkage and genome-wide association analyses. *G3 Genes Genomes Genet.* 9, 1437–1447. doi: 10.1534/g3.119.400103
- Food and Agriculture Organization (2009). “How to Feed the World in 2050,” in *Executive Summary-Proceedings of the Expert Meeting on How to Feed the World in 2050 2009 Oct 12*, (Rome: Food and Agriculture Organization).
- Frutos, E., Galindo, M. P., and Leiva, V. (2014). An interactive biplot implementation in R for modelling genotype-by-environment interaction. *Stochastic Environ. Res. Risk Assess.* 28, 1629–1641. doi: 10.1007/s00477-013-0821-z
- Fuentes-Dávila, G. (1996). “Karnal bunt,” in *Bunt and Smut Diseases of Wheat: Concepts and Methods of Disease Management*, eds R. D. Wilcoxson and E. E. Saari (Mexico: CIMMYT), 26–32. doi: 10.17221/6226-cjgpb
- Fuentes-Dávila, G., Prakash-Singh, R., Rosas-Jáuregui, I. A., Ayón-Ibarra, C. A., Félix-Valencia, P., Félix-Fuentes, J. L., et al. (2018). Evaluating advanced bread wheat lines for Karnal bunt resistance in the field during the 2012–2013 crop season. *Annu. Wheat Newsl.* 64, 25–30.
- Fuentes-Dávila, G., and Rajaram, S. (1994). Sources of resistance to *Tilletia indica* in wheat (*Triticum aestivum*). *Crop Protect.* 13, 20–24. doi: 10.1016/0261-2194(94)90131-7
- Fuentes-Dávila, G., Rajaram, S., and Singh, G. (1995). Inheritance of resistance to Karnal bunt (*Tilletia indica* Mitra) in bread wheat (*Triticum aestivum* L.). *Plant Breed.* 114, 250–252. doi: 10.1111/j.1439-0523.1995.tb00804.x
- Gaikpa, D. S., Koch, S., Fromme, F. J., Siekmann, D., Würschum, T., and Miedaner, T. (2020). Genome-wide association mapping and genomic prediction of Fusarium head blight resistance, heading stage and plant height in winter rye (*Secale cereale*). *Plant Breed.* 139, 508–520. doi: 10.1111/pbr.12810
- Gaudet, D. A., Fuentes-Dávila, G., De Pauw, R. M., and Burnett, P. A. (2001). Reactions of western Canadian spring wheat and triticale varieties to *Tilletia indica*, the causal agent of Karnal bunt. *Can. J. Plant Sci.* 81, 503–508. doi: 10.4141/p00-067
- Ginestet, C. (2011). ggplot2: elegant graphics for data analysis. *J. R. Stat. Soc. Ser. A* 174, 245–246. doi: 10.1111/j.1467-985x.2010.00676_9.x
- Gupta, V., He, X., Kumar, N., Fuentes-Dávila, G., Sharma, R. K., Dreisigacker, S., et al. (2019). Genome wide association study of Karnal bunt resistance in a wheat germplasm collection from Afghanistan. *Int. J. Mol. Sci.* 20:3124. doi: 10.3390/ijms20133124
- Gutierrez, L., Quero, G., Fernandez, S., and Brandariz, S. (2016). *lmem.Gwaser: Linear Mixed Effects Models for Genome-Wide Association Studies. R package version 0.1.0*.
- Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* 11:94. doi: 10.1186/1471-2156-11-94
- Jorgensen, J. H. (1992). Discovery, characterization, and exploitation of Mlo powdery mildew resistance in barley. *Euphytica* 63, 141–152. doi: 10.1007/bf00023919
- Joshi, L. M., Singh, D. V., Srivastava, K. D., and Wilcoxson, R. D. (1983). Karnal bunt: a minor disease that is now a threat to wheat. *Bot. Rev.* 49, 309–330. doi: 10.1007/bf02861085
- Joukhadar, R., Hollaway, G., Shi, F., Kant, S., Forrest, K., Wong, D., et al. (2020). Genome-wide association reveals a complex architecture for rust resistance in 2300 worldwide bread wheat accessions screened under various Australian conditions. *Theor. Appl. Genet.* 133, 2695–2712. doi: 10.1007/s00122-020-03626-9
- Juliana, P., Poland, J., and Huerta-Espino, J. (2019). Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nat. Genet.* 51, 1530–1539. doi: 10.1038/s41588-019-0496-6
- Kumar, M., Luthra, O. P., Yadav, N. R., Chaudhary, L., Saini, N., Kumar, R., et al. (2007). Identification of micro satellite markers on chromosomes of bread wheat showing an association with Karnal bunt resistance. *Afr. J. Biotechnol.* 6, 1617–1622.

- Kumar, S., Chawla, V., Yadav, N. R., Sharma, I., Yadav, P. K., Kumar, S., et al. (2015). Identification and validation of SSR markers for Karnal bunt (*Neovossia indica*) resistance in wheat (*Triticum aestivum*). *Indian J. Agric. Sci.* 85, 712–717.
- Kaur, M., Singh, R., Kumar, S., Mandhan, R. P., and Sharma, I. (2016). Identification of QTL conferring Karnal bunt resistance in wheat. *Indian J. Biotechnol.* 15, 34–38.
- Langridge, P., and Waugh, R. (2019). Harnessing the potential of germplasm collections. *Nat. Genet.* 51, 200–201. doi: 10.1038/s41588-018-0340-4
- Li, J., and Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95, 221–227. doi: 10.1038/sj.hdy.6800717
- Linehan, V., Thorpe, S., Andrews, N., Kim, Y., and Beaini, F. (2012). “Food demand to 2050: opportunities for Australian agriculture,” in *Paper Presented at the 42nd ABARES Outlook Conference 6-7 March 2012*, (Canberra: ACT).
- Meuwissen, T. H. E., 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
- Mitra, M. (1935). Studies on the stinking smut (bunt) of wheat with a special reference to *Tilletia indica* Mitra. *Indian J. Agric. Sci.* 5, 1–24.
- Moffat, C. S., See, P. T., and Oliver, R. P. (2015). Leaf yellowing of the wheat cultivar Mace in the absence of yellow spot disease. *Austral. Plant Pathol.* 44, 161–166. doi: 10.1007/s13313-014-0335-2
- Mujeeb-Kazi, A., Fuentes-Davilla, G., Gul, A., and Mirza, J. I. (2006). Karnal bunt resistance in synthetic hexaploid wheats (SH) derived from durum wheat \times *Aegilops tauschii* combinations and in some SH \times bread wheat derivatives. *Cereal Res. Commun.* 34, 1199–1205. doi: 10.1556/crc.34.2006.4.259
- Murray, G., and Brennan, J. (1998). The risk to Australia from *Tilletia indica*, the cause of Karnal bunt of wheat. *Austral. J. Plant Pathol.* 27, 212–225. doi: 10.1071/ap98024
- Nelson, J. C., Autrique, J. E., Fuentes-Dávila, G., and Sorrells, M. E. (1998). Chromosomal location of genes for resistance to Karnal bunt in wheat. *Crop Sci.* 38, 231–236.
- Ning, Y., Liu, W., and Wang, G. L. (2017). Balancing immunity and yield in crop plants. *Trends Plant Sci.* 22, 1069–1079. doi: 10.1016/j.tplants.2017.09.010
- Oliver, R., Lichtenzweig, J., Tan, K.-C., Waters, O., Rybak, K., Lawrence, J., et al. (2014). Absence of detectable yield penalty associated with insensitivity to *Pleosporales necrotrophic* effectors in wheat grown in the West Australian wheat belt. *Plant Pathol.* 63, 1027–1032. doi: 10.1111/ppa.12191
- Poland, J., and Rutkoski, J. (2016). Advances and challenges in genomic selection for disease resistance. *Annu. Rev. Phytopathol.* 54, 79–98. doi: 10.1146/annurev-phyto-080615-100056
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- Rice, B., and Lipka, A. E. (2019). Evaluation of RR-BLUP genomic selection models that incorporate peak genome-wide association study signals in maize and sorghum. *Plant Genome* 12, 1–4.
- Rodríguez-Álvarez, M. X., Boer, M. P., van Eeuwijk, F. A., and Eilers, P. H. (2018). Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spat. Stat.* 23, 52–71. doi: 10.1016/j.spasta.2017.10.003
- Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J. L., and Sorrells, M. (2012). Evaluation of genomic prediction methods for Fusarium head blight resistance in wheat. *Plant Genome* 5.
- Sharp, G. L., Martin, J. M., Lanning, S. P., Blake, N. K., Brey, C. W., Sivamani, E., et al. (2002). Field evaluation of transgenic and classical sources of wheat streak mosaic virus resistance. *Crop Sci.* 42, 105–110. doi: 10.2135/cropsci2002.1050
- Singh, R., Fuentes, G., Autrique, E., Huerta-Espino, J., and Singh, P. (2016). “Identification and distribution of new high yielding genotypes with resistance to Karnal bunt of wheat,” in *Proceedings of the XIX International Workshop on Smuts and Bunts*, Logan.
- Singh, S., Sehgal, D., Kumar, S., Arif, M. A., Vikram, P., Sansaloni, C. P., et al. (2020). GWAS revealed a novel resistance locus on chromosome 4D for the quarantine disease Karnal bunt in diverse wheat pre-breeding germplasm. *Sci. Rep.* 10:5999.
- Singh, S., Brown-Guedira, G. L., Grewal, T. S., Dhaliwal, H. S., Nelson, J. C., Singh, H., et al. (2003). Mapping of a resistance gene effective against Karnal bunt pathogen of wheat. *Theor. Appl. Genet.* 106, 287–292.
- Singh, S., Sharma, I., Sehgal, S. K., Bains, N. S., Guo, Z., Nelson, J. C., et al. (2007). Molecular mapping of QTLs for Karnal bunt resistance in two recombinant inbred populations of bread wheat. *Theor. Appl. Genet.* 116, 147–154. doi: 10.1007/s00122-007-0654-6
- Singh, S., Vargas, M., Crossa, J., Singh, P. K., Bains, N. S., Singh, K., et al. (2012). Multi trait and multi-environment QTL analyses for resistance to wheat diseases. *PLoS One* 7:e38008. doi: 10.1371/journal.pone.0038008
- Slavov, G. T., Nipper, R., Robson, P., Farrar, K., Allison, G. G., Bosch, M., et al. (2014). Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass *Miscanthus sinensis*. *New Phytol.* 201, 1227–1239. doi: 10.1111/nph.12621
- Stekhoven, D. J., and Bühlmann, P. (2012). MissForest - non-parametric missing value imputation for mixed-type data. *Bioinformatics* 28, 112–118. doi: 10.1093/bioinformatics/btr597
- Tan, M.-K., El Bouhssini, M., Emebiri, L. C., Wildman, O., Tadesse, W., and Ogbonnaya, F. C. (2015). A SNP marker for the selection of HfrDrd, a Hessian fly-response gene in wheat. *Mol. Breed.* 35:216.
- Vocke, G., Allen, E., and Price, J. M. (2010). *The Economic Impact of Karnal Bunt Phytosanitary Wheat Export Certificates*. Outlook Report No. WHS-10h-01. Washington, DC: US Department of Agriculture, Economic Research Service.
- Vyska, M., Cuniffe, N., and Gilligan, C. (2016). Trade-off between disease resistance and crop yield: a landscape-scale mathematical modelling perspective. *J. R. Soc. Interf.* 13:20160451. doi: 10.1098/rsif.2016.0451
- Warham, E. J. (1986). Karnal bunt disease of wheat: a literature review. *Trop. Pest Manag.* 32, 229–242. doi: 10.1016/b978-0-12-819527-7.00015-7
- White, B., Day, C., Christopher, M., and van Klinken, R. (2016). *Should we Invest Now in Cereal Pre-Breeding for Biosecurity Threats?* Working Paper 1605, School of Agricultural and Resource Economics, University of Western Australia, Crawley, Australia. Available online at: <https://ideas.repec.org/p/ags/uwauwp/236736.html> (assessed 12 June, 2021).
- 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., 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
- Zhao, Y., Mette, M. F., Gowda, M., Longin, C. F., and Reif, J. C. (2014). Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity* 112, 638–645. doi: 10.1038/hdy.2014.1

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.

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 Emebiri, Hildebrand, Tan, Juliana, Singh, Fuentes-Davila 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.



Breeding With Major and Minor Genes: Genomic Selection for Quantitative Disease Resistance

Lance F. Merrick¹, Adrienne B. Burke¹, Xianming Chen² and Arron H. Carter^{1*}

¹ Department of Crop and Soil Sciences, Washington State University, Pullman, WA, United States, ² United States Department of Agriculture-Agricultural Research Service Wheat Health, Genetics and Quality Research Unit, Department of Plant Pathology, Washington State University, Pullman, WA, United States

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Eric Dinglasan,
The University of
Queensland, Australia
Colin Hiebert,
Agriculture and Agri-Food
Canada, Canada

Jemanesh Kifetew Haile,

University of Saskatchewan, Canada

*Correspondence:

Arron H. Carter
ahcarter@wsu.edu

Specialty section:

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

Received: 23 May 2021

Accepted: 08 July 2021

Published: 06 August 2021

Citation:

Merrick LF, Burke AB, Chen X and
Carter AH (2021) Breeding With Major
and Minor Genes: Genomic Selection
for Quantitative Disease Resistance.
Front. Plant Sci. 12:713667.
doi: 10.3389/fpls.2021.713667

Disease resistance in plants is mostly quantitative, with both major and minor genes controlling resistance. This research aimed to optimize genomic selection (GS) models for use in breeding programs that are needed to select both major and minor genes for resistance. In this study, stripe rust (*Puccinia striiformis* Westend. f. sp. *tritici* Erikss.) of wheat (*Triticum aestivum* L.) was used as a model for quantitative disease resistance. The quantitative nature of stripe rust is usually phenotyped with two disease traits, infection type (IT) and disease severity (SEV). We compared two types of training populations composed of 2,630 breeding lines (BLs) phenotyped in single-plot trials from 4 years (2016–2020) and 475 diversity panel (DP) lines from 4 years (2013–2016), both across two locations. We also compared the accuracy of models using four different major gene markers and genome-wide association study (GWAS) markers as fixed effects. The prediction models used 31,975 markers that are replicated 50 times using a 5-fold cross-validation. We then compared GS models using a marker-assisted selection (MAS) to compare the prediction accuracy of the markers alone and in combination. GS models had higher accuracies than MAS and reached an accuracy of 0.72 for disease SEV. The major gene and GWAS markers had only a small to nil increase in the prediction accuracy more than the base GS model, with the highest accuracy increase of 0.03 for the major markers and 0.06 for the GWAS markers. There was a statistical increase in the accuracy using the disease SEV trait, BLs, population type, and combining years. There was also a statistical increase in the accuracy using the major markers in the validation sets as the mean accuracy decreased. The inclusion of fixed effects in low prediction scenarios increased the accuracy up to 0.06 for GS models using significant GWAS markers. Our results indicate that GS can accurately predict quantitative disease resistance in the presence of major and minor genes.

Keywords: genomic selection, fixed-effect, disease resistance, stripe rust, genome-wide associate studies, rrBLUP

INTRODUCTION

Plant breeding programs select and improve both qualitative and quantitative traits. Qualitative traits are controlled by a few large-effect genes that are readily detectable and follow a Mendelian inheritance (Chen, 2013). In contrast, quantitative traits are controlled by several small-effect genes that are difficult to distinguish and controlled by quantitative trait loci (QTL; Bernardo, 2008). The genetic control of a trait determines the types of selection that will be most effective for improvement. However, disease resistance can be either a qualitative or a quantitative trait, and, therefore, the most effective method of improvement varies (Poland and Rutkoski, 2016). Breeding for disease resistance is a major goal for most breeding programs due to the effect of the disease on yield and quality performance.

Breeding for qualitative disease resistance is controlled by one or two large-effect alleles, called resistance (R) genes and further referred to as major genes (Agrios, 2005). Qualitative disease resistance generally follows a race-specific resistance and quickly degrades due to the rapid evolution of new pathogen races (Chen, 2005). Major gene pyramiding can reduce the possibility of major genes by combining multiple major genes to provide a more durable resistance to multiple pathogen races into a single line. Pyramiding is implemented through a marker-assisted selection (MAS) and has been an effective method for various crops (Wang et al., 2001, 2017; Pietrusińska et al., 2011; Bai et al., 2012; Jiang et al., 2012; Liu et al., 2016b; Singh et al., 2017). Successful implementation of major genes relies on identifying the useful sources of the genes, finding the linked markers, confirming the effect in different genetic backgrounds, and finally, deploying said major genes (Bernardo, 2008). Major gene implementation is further complicated when it comes to selecting multiple major genes simultaneously for gene pyramiding. A large population is needed to screen and select the lines with more than one gene in early generations while still maintaining enough lines to select for other traits in later generations (Poland and Rutkoski, 2016). The difficulty can be further attributed to unfavorable linkage and multiple major gene sources (Bernardo, 2008).

Breeding for quantitative resistance conferred by minor-effect genes or a combination of minor and major genes tends to produce a more durable resistance in breeding lines (BLs) because it relies on multi-resistant alleles. Breeding for quantitative resistance requires multiple breeding cycles to improve resistance gradually (Poland and Rutkoski, 2016). The breeding method for quantitative resistance is similar to the methodology used for other complex traits such as grain yield (Rutkoski et al., 2014; Poland and Rutkoski, 2016; González-Camacho et al., 2018). Similar to qualitative resistance, selecting for quantitative resistance can be completed throughout the breeding process, but disease resistance is commonly completed in earlier generations to select for other traits further in the program. Therefore, selecting for quantitative resistance in earlier generations can be difficult due to the lack of replication and environments. However, selecting for resistance in later generations reduces genetic gain due to the selection for other traits (Poland and Rutkoski, 2016). Both methods, therefore, reduce the effectiveness of breeding quantitative resistance.

One such trait that displays both qualitative and quantitative resistance is stripe rust, also called yellow rust (Yr), caused by *Puccinia striiformis* Westend. f. sp. *tritici* Erikss.

Stripe rust is one of the most devastating diseases of wheat (*Triticum aestivum* L.) and is highly destructive in the western USA (Chen, 2005; González-Camacho et al., 2018; Liu et al., 2019). Stripe rust can cause more than 90% yield losses in fields planted with susceptible cultivars (Liu et al., 2020). The use of resistance varieties and the applications of fungicide are the primary methods to control stripe rust (Chen and Line, 1995; Liu et al., 2020). Stripe rust resistance is categorized into qualitative all-stage resistance (ASR) and quantitative adult-plant resistance (APR).

All-stage resistance is conferred by race-specific genes that are inherited qualitatively with a life span of ~3.5 years per gene (Case et al., 2014; Chen and Kang, 2017). There are more than 300 identified QTL conferring resistance to stripe rust (Wang and Chen, 2017). The identification of a large number of major genes shows numerous resistance alleles available for breeding purposes in various varieties and populations. Previously, major genes *Yr5* and *Yr15* have been shown to be effective against all races of the stripe rust pathogen in the USA (Wang and Chen, 2017). However, virulence to *Yr5* has been demonstrated in a few countries not including the USA (Wellings et al., 2009; Zhang et al., 2020; Kharouf et al., 2021; Tekin et al., 2021). Virulence to *Yr15* has only been documented in Afghanistan (Gerechter-Amitai et al., 1989). The virulence to these genes demonstrates the need to not rely on any single major gene to provide resistance in a cultivar.

Adult-plant resistance is usually a non-race-specific quantitative resistance that is associated with durable resistance with some genes being effective for more than 60 years (Chen, 2013). APR is often affected by temperature and also can be referred to as high-temperature adult-plant (HTAP) resistance, which is often controlled by more than one gene mainly with additive effect (Chen and Line, 1995; Chen et al., 1995; Liu et al., 2019). HTAP resistance is influenced by the temperature and age of the plants. As the temperature increases, the plant becomes more resistant, and rust development slows down (Chen, 2005). However, to confirm HTAP, greenhouse studies with different temperature ranges need to be conducted (Chen, 2005). HTAP resistance and APR are conferred by different loci with varying effects and often display partial resistance, making them difficult to incorporate into new cultivars (Chen and Line, 1995; Liu et al., 2019). Consequently, APR or HTAP resistance must be improved over multiple selection cycles as mentioned previously (Rutkoski et al., 2014; Poland and Rutkoski, 2016; González-Camacho et al., 2018). APR is generally expressed in the later stages of wheat, whereas ASR is expressed throughout the lifecycle of the plant (Wang and Chen, 2017). Therefore, it is difficult to identify APR genes due to the masking of their effect by ASR genes. The masking of ASR genes and the quantitative nature of APR genes result in much of the APR resistance in a population being uncharacterized. It is recommended to combine both ASR and APR genes to take advantage of both types of resistance limitations (Wang and Chen, 2017). The lack of ASR durability coupled with the challenge in identifying and breeding APR

creates a unique opportunity for genomic selection (GS). In addition, major ASR genes are known to interact with APR and including them in GS models as fixed effect have increased prediction accuracy (Bernardo, 2014; Rutkoski et al., 2014; Arruda et al., 2016).

In many crops, the difficulty in selecting for qualitative and quantitative disease resistance (similar to stripe rust) creates an opportunity for GS to integrate quantitative resistance by accounting for small-effect alleles in the presence of large-effect major genes without the development and analysis of mapping populations and techniques (Poland and Rutkoski, 2016). The goal of this study was to determine the most accurate GS method to select for disease resistance in the presence of both major and minor genes. Wheat stripe rust was used as an example as most plant breeders try to capture the additive effects of both ASR and APR simultaneously. The identified GS approaches will be a valuable tool for breeders to facilitate cultivar and parental selection for accumulating favorable alleles for disease resistance in the presence of major and minor resistance genes (Rutkoski et al., 2014; Michel et al., 2017).

MATERIALS AND METHODS

Phenotypic Data

Two training populations were used to compare the inclusion of fixed-effect markers in populations with different frequencies of stripe rust genes. The first training population consists of F_{3:5} and double-haploid soft white winter wheat BLs developed by the Washington State University (WSU) winter wheat breeding program. The BL population was evaluated for stripe rust in the unreplicated single-plot trials in Pullman and Lind, Washington planted in 2016, 2017, 2018, and 2020 growing seasons (Table 1). Due to the unreplicated nature of the single-plot trails, each trial consisted of unique lines, which resulted in a total of 2,630 lines for all years and locations. The year 2019 was not included due to the lack of adequate disease SEV in our trials. The BL population was previously selected for stripe rust resistance in headrow plots the year previous to unreplicated trials. Susceptible BLs in headrow plots were culled and not included in the BL population, which represents a prior selected, closely related BL population with similar pedigree sources of stripe rust resistance. The second training population consists of diverse association mapping panel [diversity panel (DP)] trials evaluated in unreplicated trials in Central Ferry and Pullman, Washington from 2013 to 2016 with the same 475 lines represented in each trial (Table 1). The mapping panel consists of varieties and BLs from at least six soft white winter wheat breeding programs in the Pacific Northwest (PNW) and represents diverse backgrounds with the potential sources of stripe rust resistance.

The measured disease traits were stripe rust IT and SEV. The recordings of these traits were dependent on natural infection and stripe rust incidence at the time of observation and were not previously inoculated. Some trials had three observations for stripe rust and were identified with sequential numbers. The first recording was taken soon after the emergence of a flag leaf, the second was taken again after anthesis, and the third was taken in the early milk stage. The trials with only one observation were

recorded right after anthesis for responses in the adult plant stage as stripe rust was not present in the field during earlier growth stages. IT was recorded based on a 0–9 scale (Line and Qayoum, 1992). SEV was recorded as the percentage of the leaf-infected area using the modified Cobb Scale (Peterson et al., 1948). Table 1 summarizes environments, years, genotyped individuals, and the measurements taken for each trial during which stripe rust was recorded. However, due to the nature of APR being effective in the adult stage and the fact that not all trials had multiple recordings, only the last observation for each trial was used to measure the disease traits for APR.

To account for differences in disease pressure under different environments, a two-step adjusted mean method by which a linear model was implemented to adjust both IT and SEV means within and across environments was used. Then, a mixed linear model was used to calculate genomic estimated breeding values (GEBVs; Ward et al., 2019). Means from the stripe rust data collected in the unreplicated trials were adjusted using the residuals calculated for the unreplicated genotypes in individual environments and across environments using the modified augmented complete block design (ACBD) model (Federer, 1956; Goldman, 2019). The adjustments were made according to the method implemented in Merrick and Carter (2021), with the full model across environments as follows:

$$Y_{ij} = \text{Block}_i + \text{Check}_j + \text{Env}_k + \text{Block}_i \times \text{Env}_k + \text{Check}_j \times \text{Env}_k + \varepsilon_{ijk} \quad (1)$$

where Y_{ij} is the trait of interest, either IT or SEV; Block_i is the fixed effect of the i th block and k th trial; Check_j is the fixed effect of the j th replicated check cultivar; Env_k is the fixed effect of the k th trial; and ε_{ijk} denote the residual errors.

Heritability on a genotype-difference basis for broad-sense heritability was calculated using the variance components from the models implemented in Merrick and Carter (2021) and using the best linear unbiased predictors for both individual environments and across environments with the formula:

$$H^2_{\text{Cullis}} = 1 - \frac{\bar{v}_{\Delta}^{\text{BLUP}}}{2\sigma_g^2} \quad (2)$$

where σ_g^2 and $\bar{v}_{\Delta}^{\text{BLUP}}$ are the genotype variance and mean variance of the BLUPs, respectively (Cullis et al., 2006). In general, the broad-sense heritability measurement is not suitable for an unreplicated, unbalanced multi-environment trial, and, therefore, narrow-sense heritability was not calculated (Schmidt et al., 2019).

Genotypic Data

Lines were genotyped using genotyping-by-sequencing (GBS; Elshire et al., 2011) through the North Carolina State Genomics Sciences Laboratory in Raleigh, NC, USA, using the restriction enzymes *MspI* and *PstI* (Poland et al., 2012). Genomic DNA was isolated from seedlings in the one-leaf to three-leaf stage using Qiagen BioSprint 96 Plant kits and the Qiagen BioSprint 96 workstation (Qiagen, Germantown, MD, USA). DNA libraries were prepared following the protocol of DNA digestion with *PstI*

TABLE 1 | Training populations for stripe rust IT and SEV in Central Ferry, Lind, and Pullman, WA, USA from 2013 to 2020.

| Population ^a | Year | Trials | Locations | Lines ^b | IT 1 ^c | SEV 1 ^d | IT 2 | SEV 2 | IT 3 | SEV 3 |
|-------------------------|-----------|--------|-----------|--------------------|-------------------|--------------------|------|-------|------|-------|
| DP | 2013 | 2 | 2 | 475 | X | X | X | X | X | X |
| DP | 2014 | 2 | 2 | 474 | X | X | X | X | X | X |
| DP | 2015 | 2 | 2 | 474 | X | X | X | X | X | X |
| DP | 2016 | 1 | 1 | 474 | X | X | X | X | X | X |
| DP | 2013–2014 | 4 | 4 | 475 | X | X | X | X | X | X |
| DP | 2013–2015 | 6 | 6 | 475 | X | X | X | X | X | X |
| DP | 2013–2016 | 7 | 7 | 475 | X | X | X | X | X | X |
| BL | 2016 | 2 | 2 | 304 | X | X | X | | | |
| BL | 2017 | 4 | 2 | 728 | X | X | X | X | X | X |
| BL | 2018 | 3 | 2 | 1,239 | X | X | X | X | X | X |
| BL | 2020 | 1 | 1 | 373 | X | X | | | | |
| BL | 2016–2017 | 6 | 4 | 1,029 | X | X | X | X | X | X |
| BL | 2016–2018 | 9 | 6 | 2,262 | X | X | X | X | X | X |
| BL | 2016–2020 | 10 | 7 | 2,630 | X | X | X | X | X | X |

^aDP, Diversity panel; BLs, Breeding lines.

^bLines, Unique lines in the training population.

^cIT, Infection type.

^dSEV, Disease severity.

X, Indicates measurement recorded.

and *MspI* restriction enzymes (Poland et al., 2012). Genotyping-by-sequencing (GBS; Elshire et al., 2011) was conducted at North Carolina State University Genomic Sciences Laboratory with either an Illumina HiSeq 2500 or a NovaSeq 6000. DNA library barcode adapters, DNA library analysis, and sequence single-nucleotide polymorphism (SNP) calling were provided by the USDA Eastern Regional Small Grains Genotyping Laboratory (Raleigh, NC, USA). Sequences were aligned to the Chinese Spring International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 (Appels et al., 2018), using the Burrows–Wheeler Aligner (BWA) 0.7.17 (Li and Durbin, 2009). Genetic markers with more than 20% missing data, a minor allelic frequency of <5%, and the markers that were monomorphic were removed. Markers were then imputed using Beagle version 5.0 and filtered once more for markers less than a 5% minor allelic frequency (Browning et al., 2018). A total of 31,975 SNP markers for the 475 unique DP lines and 2,630 BLs were obtained from GBS. Principal components for the markers were calculated using the function “prcomp,” and a biplot with *k*-mean clusters was created using the function “autoplot” in R (R Core Team, 2018). Cluster number for *k*-means was calculated according to the elbow method using a screen plot with the identification of the optimal number of clusters when the total intracluster variation was minimized.

Major rust-resistant genes observed to be common in the WSU breeding population are *Yr10*, *Yr17*, *Lr68*, and *Qyr.wpg-1B.1*, and molecular marker data for these genes were included as fixed effects in our GS models. All winter wheat lines were genotyped using Kompetitive Allele Specific PCR (KASP®) assay for *Yr17*, *Lr68*, and *Qyr.wpg-1B.1* in the WSU winter wheat breeding laboratory. The *Yr17* gene (Helguera et al., 2003) was screened using the KASP marker developed by Milus et al. (2015). The *Lr68* leaf rust resistance gene (Herrera-Foessel et al., 2012)

was screened using the KASP marker developed by Rasheed et al. (2016). Although leaf rust resistance is not commonly selected in the US PNW breeding programs, this gene was found in a large proportion of BLs, and thus was hypothesized that it might have been selected congruently with stripe rust resistance. The APR QTL *Qyr.wpg-1B.1* reported on chromosome 1B by Naruoka et al. (2015) was screened using the marker *IWB12603* (Mu et al., 2020). The KASP assays were performed using PACE™ Genotyping Master Mix (3CR Bioscience, Essex, UK) following the instructions of the manufacturer, and endpoint genotyping was conducted on fluorescence using a Lightcycler 480 Instrument II (Roche, Indianapolis, IN, USA). The previously reported ASR gene *Yr10* (Frick et al., 1998) was screened with a microsatellite marker *Xpsp3000* developed by Bariana et al. (2002). The microsatellite marker *Xpsp3000* was run using PCR products, which were separated on an ABI3730XL DNA fragment analyzer (Applied Biosystems, Waltham, MA, USA), and alleles were scored with the GeneMarkerv4.0 software (SoftGenetics, State College, PA, USA), in collaboration with the USDA Western Regional Small Grains Genotyping Laboratory in Pullman, Washington.

Genome-Wide Association Model

In addition to the inclusion of molecular markers for major rust-resistant genes as fixed effects, the markers identified through genome-wide association studies (GWASs) were included through *de novo* GWAS. This method is further referred to as GWAS-assisted GS (GWAS-GS). The GWAS-GS was implemented according to McGowan et al. (2020). Briefly, a proper cross-validation using GWAS was conducted using BLINK in the genome association and prediction integrated tool (Liu et al., 2016a; Tang et al., 2016; Huang et al., 2019; GAPIT) with three principal components fitted as fixed effects on the

training population only. Three principal components were used because they were previously observed to be most reliable in accounting for a population structure for yield and agronomic traits in winter wheat for the same populations (Lozada et al., 2017). In accordance to advice put forward by Rice and Lipka (2019), the first method of GWAS-GS included only significant markers based on a Bonferroni cutoff of 0.05 (GWAS_B). Due to our cross-validation scheme, different significant markers for GWAS_B were identified in each cross-validation, year, and population. Therefore, significant markers were not presented. For the remaining GWAS-GS methods, the markers were ordered by the degree of statistical significance based on the values of p from the smallest to largest. We compared the inclusion of the top 5, 10, 25, 50, and 100 most significant markers as fixed effects (GWAS_5, GWAS_10, GWAS_25, GWAS_50, and GWAS_100).

Prediction Models

Marker-Assisted Selection Model

Single and multiple regression models were used as MAS models to compare major rust-resistant markers and the predictive ability of *de novo* GWAS markers alone and in combination. The fixed-effect multiple regression model is described as follows:

$$y_i = \mu + X_1\beta_1 + \dots X_i\beta_i + \varepsilon_i \quad (3)$$

where y_i is the observed phenotypic value of the i th individual, μ is the mean, X_i is the genotype of the marker i , β_i is the effect of the i th marker, and ε_i is the residual error term.

GS Model

rrBLUP was used as the base GS model and was implemented using the package “rrBLUP” (Endelman, 2011). rrBLUP was used as the base model due to the nonplacement of the ridge regression penalty implemented by rrBLUP on the fixed effects, allowing a large effect on the model. Further, rrBLUP has shown to outperform other models when integrating fixed effects into the models and in predicting disease resistance (Rutkoski et al., 2014; Arruda et al., 2016; Poland and Rutkoski, 2016; Muleta et al., 2017). The basic rrBLUP model is described as follows (Rice and Lipka, 2019):

$$y_i = \mu + \sum_{k=1}^p x_{ik}\beta_k + \varepsilon_i \quad (4)$$

where y_i is the observed phenotypic value of the i th individual, μ is the mean, x_{ik} is the genotype of the k th marker and i th individual, p is the total number of markers, β_k is the estimated random marker effect of the k th marker, and ε_i is the residual error term.

GS Model With Fixed Effects

To evaluate the effect of major and *de novo* GWAS markers on the prediction accuracy of GS models, we used the rrBLUP model as described (Rice and Lipka, 2019):

$$y_i = \mu + \sum_{j=1}^m x_{ij}\alpha_j + \sum_{k=1}^p x_{ik}\beta_k + \varepsilon_i \quad (5)$$

where y_i is the observed phenotypic value of the i th individual, μ is the mean, x_{ij} is the j th marker of the i th individual, m is the number of markers included as fixed-effect covariates, α_j is the fixed additive effect of the j th marker, x_{ik} is the genotype of the k th marker and i th individual, p is the total number of markers, β_k is the estimated random marker effect of the k th marker, and ε_i is the residual error term.

Prediction Accuracy and Schemes

The prediction accuracy for the GS was reported using Pearson correlation coefficients, and a prediction bias was reported using a root mean square error (RMSE) between GEBVs and their respective adjusted means using the function “cor” in R (R Core Team, 2018). The effect of fixed-effect markers on prediction accuracy was assessed using a 5-fold cross-validation scheme and independent validation sets for IT and SEV in the DP and BL training populations. The two populations were used to compare the effects of the significant markers in populations with different genetic relatedness, frequency of markers, and sources of resistant pedigrees. GS models were conducted with 5-fold cross-validation by including 80% of the samples in the training population and predicting the GEBVs of the remaining 20% (Lozada and Carter, 2019). One replicate consists of the five model iterations, where the population is split into five different groups. This was completed 50 times. As mentioned previously for the GWAS-GS, the GWAS was conducted on 80% of the lines, and then the markers are included in the GS model to predict the remaining 20% of the lines. Independent validation sets were then performed on a yearly basis by combining the two training populations and environments together per year. This allows the evaluation of models in a realistic breeding situation in which we combine all available data to build a training population.

The training populations were evaluated for cross-validations on a yearly basis and over combined years and trials. We assessed each year independently using cross-validations. We then created prediction models starting with the earliest trial and then a new model with the addition of each subsequent trial to evaluate a genotype-by-environment interaction, continuous training of a prediction model, and the effect of different races of *P. striiformis* f. sp. *tritici*. Independent validation sets were first conducted using continuous training. For example, the earliest year, i.e., 2013, was used to predict the following year, i.e., 2014. The years were then combined to predict the following year, i.e., 2013 and 2014 to predict 2015, and this process was continued until the years 2013–2018 were used to predict 2020. Using this scenario, the first 3 years, 2013–2015, consisted of the DP lines alone, and therefore, each year consisted of the same lines. With the inclusion of the years 2016–2020, unique lines from the BL were added each year due to the fact that each trial in the BL consisted of unique lines only phenotyped in a single trial as mentioned previously.

All GS and MAS models and scenarios were analyzed using WSU’s Kamiak high performance computing cluster (Kamiak, 2021). Model and year comparisons were evaluated by using a Tukey’s honestly significant difference (HSD) test implemented in the “agricolae” package in R (R Core Team, 2018; de Mendiburu and de Mendiburu, 2019). The comparison of models

was then plotted for a visual comparison using “ggplot2” in R (Wickham, 2011; R Core Team, 2018).

RESULTS

Phenotypic Data

Stripe rust phenotyping was dependent on natural infection. Therefore, it is important to evaluate GS models in different years to account for environments with little to no variation in stripe rust SEV and pathogen race changes. Overall, the maximum IT and SEV were relatively high for each scale, indicating the presence of adequate stripe rust SEV in each trial (**Table 2**). The BL had relatively high coefficient of variations (CVs) for each trial. However, the heritability was very high, ranging from 0.60 to 0.96 across traits and trials, indicating adequate screening trials for stripe rust. Further, the phenotypic correlations between IT and SEV were relatively high in the DP, ranging from 0.67 in 2013 to 0.88 in 2015 (**Table 3**). The phenotypic correlation in the BL between IT and SEV was similarly high, ranging from 0.70 in 2016 to 0.86 in 2018.

The inclusion of multiple environments creates a challenge for GS models due to the genotype–environment interaction (GEI). There were significant differences between the majority of years for each population and trait (**Figures 1A–D**). The ranges for both IT and SEV were large, indicating both resistant and nonresistant varieties within the populations. The mean IT and SEV were also lower in the BL compared to the DP (**Figures 1A–D**; **Table 2**). In comparison to the DP, the BL population consisted of a larger proportion of resistant cultivars, which was expected as these had previously been selected under field conditions. SEV displayed a large concentration of values near zero, specifically in the year 2018 (**Figure 1D**). Significant differences of each year indicate an environmental effect that needs to be accounted for within the prediction models.

In addition to GEI, stripe rust races may change from year to year, which creates an opportunity for major genes to be overcome by virulent races. The USDA stripe rust lab records race frequencies and virulence each year (<https://striperust.wsu.edu/races/data/>). The major stripe rust races for each year was either PSTv-37 or PSTv-52 with the exception of 2017 and 2020, which had large frequencies for PSTv-37 (**Supplementary Table 1**). The other races with higher frequencies included PSTv-39, PSTv-322, PSTv-48, PSTv-79, PSTv-11, and PSTv-73. Therefore, the difference in race change was not a major factor in prediction scenarios.

Genotypic Data

The major rust genes present in the WSU winter wheat breeding program germplasm are *Yr10*, *Yr17*, *Lr68*, and *Qyr.wpg-1B*. The frequency of genotypes, as determined by the previously described molecular markers for each of these genes, is presented in **Table 4**. Similar frequencies in both populations were observed for the homozygous resistant allele for *Lr68* with 50 and 46% in the DP and BL, respectively. The frequency of the marker for *Yr10* and *IWB12603* was much higher in the DP than in the BL with *Yr10* having a relatively high frequency of 53% in the DP. However, the homozygous resistant allele for *Yr17* was much

higher in the BL (38%) than in the DP (19%). There was also a wide combination of homozygous resistant alleles within each population (**Figure 2**).

The principal component biplot using the GBS SNP markers over the combined DP and BL training populations accounted for only 9.1% of the total genetic variation, indicating a large population structure (**Figure 3**). PC1 explained 5.4% of the variation, and PC2 explained 3.7% of the variation. The biplot revealed three main clusters over the combined populations using *k*-means clustering. A majority of lines in both the DP and BL were included in the first cluster with 355 and 2,107 lines, respectively. The second cluster also displayed a mixture of DP lines and BLs with 108 and 219 lines, respectively. Finally, the third cluster included mainly BLs with 12 lines in the DP and 304 lines from the BL.

Cross-Validations

Major Markers

Multiple comparisons for accuracy and RMSE between the inclusion of each major molecular marker for the known rust-resistant genes individually and in combination (ALL_M) were completed for both populations in and across the years for IT and SEV (**Supplementary Tables 2, 3**). The markers for major rust genes were included as fixed effects and compared to the base rrBLUP model and MAS models with the markers as variables alone (**Figures 4, 5**; **Supplementary Tables 4, 5**). Within individual years in the BL, the rrBLUP base model reached a high accuracy of 0.65 in 2018 and 2020 for IT and 0.68 in 2018 for SEV. The effects of the major markers varied from year to year, but the marker for *Yr17* showed an increase in the prediction accuracy for every year except in 2018 for IT (**Supplementary Table 2**) and in 2018 and across 2016–2018 and 2016–2020 for SEV (**Supplementary Table 3**). A majority of markers had relatively low prediction accuracies for MAS with the exception of the *Yr17* marker that reached an accuracy of 0.05 and 0.42 for IT and SEV, respectively (**Supplementary Tables 4, 5**). When all markers were combined, similar accuracies were attained and compared to the time of inclusion of only the *Yr17* marker in both the rrBLUP model and MAS models. The remainder of major rust gene markers with the exception of *Lr68* increased the accuracy within specific years but had less consistency than the *Yr17* marker for. The largest differences from the rrBLUP model within a single year in the BL were seen in 2016 for GS models (**Supplementary Figure 1**). In 2016, the combination of both *Yr10* and *Yr17* markers increased the accuracy by 0.06 for SEV. *Yr17* and the combination of markers only slightly increased the accuracy across environments with an increase in IT of 0.01 (**Supplementary Figure 2**). Additionally, the RMSE was similar between all markers and rrBLUP for both traits (**Supplementary Figure 3**; **Supplementary Tables 2, 3**). However, for MAS, RMSE was higher than all GS models. In MAS models, the RMSE was lower for *Yr17*, and ALL_M, compared to another marker, with SEV having a much higher error than IT for the majority of years. The individual years of 2018 and 2020 displayed a higher RMSE compared to the other individual years and combined years (**Supplementary Figure 3**; **Supplementary Tables 4, 5**).

TABLE 2 | Stripe rust IT and disease SEV heritability (H^2) and trial statistics for unadjusted phenotypes in the DP and BL training population phenotypes from 2013 to 2016 and 2016 to 2020.

| Population | Year | Trait | H^2 | CV ^a | Max ^b | Mean | Min ^c | SD ^d |
|------------|-----------|-------|-------|-----------------|------------------|------|------------------|-----------------|
| DP | 2013 | IT | 0.85 | 52.31 | 9 | 3 | 1 | 2 |
| DP | 2014 | IT | 0.82 | 57.13 | 9 | 4 | 1 | 2 |
| DP | 2015 | IT | 0.89 | 43.82 | 9 | 5 | 1 | 2 |
| DP | 2016 | IT | 0.84 | 46.23 | 9 | 4 | 0 | 2 |
| DP | 2013–2014 | IT | 0.93 | 55.58 | 9 | 4 | 1 | 2 |
| DP | 2013–2015 | IT | 0.94 | 52.66 | 9 | 4 | 1 | 2 |
| DP | 2013–2016 | IT | 0.95 | 51.77 | 9 | 4 | 0 | 2 |
| DP | 2013 | SEV | 0.91 | 108.31 | 100 | 22 | 2 | 24 |
| DP | 2014 | SEV | 0.78 | 116.04 | 90 | 24 | 2 | 28 |
| DP | 2015 | SEV | 0.92 | 72.78 | 100 | 43 | 2 | 32 |
| DP | 2016 | SEV | 0.89 | 70.57 | 100 | 36 | 0 | 26 |
| DP | 2013–2014 | SEV | 0.93 | 112.77 | 100 | 23 | 2 | 26 |
| DP | 2013–2015 | SEV | 0.96 | 99.31 | 100 | 30 | 2 | 30 |
| DP | 2013–2016 | SEV | 0.96 | 94.90 | 100 | 31 | 0 | 29 |
| BL | 2016 | IT | 0.90 | 87.56 | 8 | 3 | 0 | 2 |
| BL | 2017 | IT | 0.83 | 83.36 | 9 | 3 | 0 | 2 |
| BL | 2018 | IT | 0.79 | 172.49 | 8 | 2 | 0 | 3 |
| BL | 2020 | IT | 0.96 | 93.03 | 8 | 3 | 0 | 3 |
| BL | 2016–2017 | IT | 0.83 | 84.06 | 9 | 3 | 0 | 2 |
| BL | 2016–2018 | IT | 0.79 | 115.26 | 9 | 2 | 0 | 3 |
| BL | 2016–2020 | IT | 0.80 | 113.22 | 9 | 2 | 0 | 3 |
| BL | 2016 | SEV | 0.86 | 152.31 | 80 | 9 | 0 | 13 |
| BL | 2017 | SEV | 0.88 | 131.79 | 90 | 16 | 0 | 21 |
| BL | 2018 | SEV | 0.60 | 212.74 | 80 | 11 | 0 | 24 |
| BL | 2020 | SEV | 0.96 | 125.06 | 80 | 18 | 0 | 23 |
| BL | 2016–2017 | SEV | 0.89 | 136.36 | 90 | 15 | 0 | 20 |
| BL | 2016–2018 | SEV | 0.85 | 165.25 | 90 | 13 | 0 | 22 |
| BL | 2016–2020 | SEV | 0.86 | 160.47 | 90 | 14 | 0 | 22 |

^aCV, Coefficient of variation.^bMax, Maximum.^cMin, Minimum.^dSD, Standard deviation.**TABLE 3 |** Phenotypic correlations between IT and disease SEV.

| Population | Year 1 | Year 2 | Year 3 | Year 4 | Year 1–2 | Year 1–3 | Year 1–4 |
|------------|--------|--------|--------|--------|----------|----------|----------|
| DP | 0.79 | 0.67 | 0.88 | 0.82 | 0.76 | 0.85 | 0.86 |
| BL | 0.70 | 0.80 | 0.86 | 0.85 | 0.76 | 0.83 | 0.83 |

Phenotypic correlations between IT and disease SEV for Pacific Northwest (PNW) winter wheat within both the diversity panel (DP) lines and breeding lines (BL) phenotyped from 2013 to 2020 in Central Ferry, Lind, and Pullman, WA, USA.

Within individual years in the DP, the rrBLUP base model reached an accuracy of 0.55 for IT (**Supplementary Table 2**) and 0.64 for SEV in 2013 (**Supplementary Table 3**). Across years, IT reached 0.56 in 2013–2016 (**Supplementary Table 2**) and SEV reached 0.69 in 2013–2014 (**Supplementary Table 3**). In the DP, the major rust markers had a less effect on the prediction accuracy, with *Yr10* being the only marker that increased the accuracy from the base rrBLUP model and at a maximum of 0.01. For MAS, the combination of markers resulted

in the least reduction of accuracy with a maximum reduction of 0.10 in 2015 for IT (**Supplementary Table 4**) and 0.07 in 2016 for SEV (**Supplementary Table 5**). Markers for *Yr10* and *IWB12603* also had the largest effect on MAS models. The largest differences from the rrBLUP model within a single year in the DP were seen in 2015 for GS models (**Supplementary Figure 4**). In 2015, the *Yr10* marker increased the accuracy by 0.01. There were no increases in the accuracy across any combination of environments in the DP. The results for RMSE were similar to

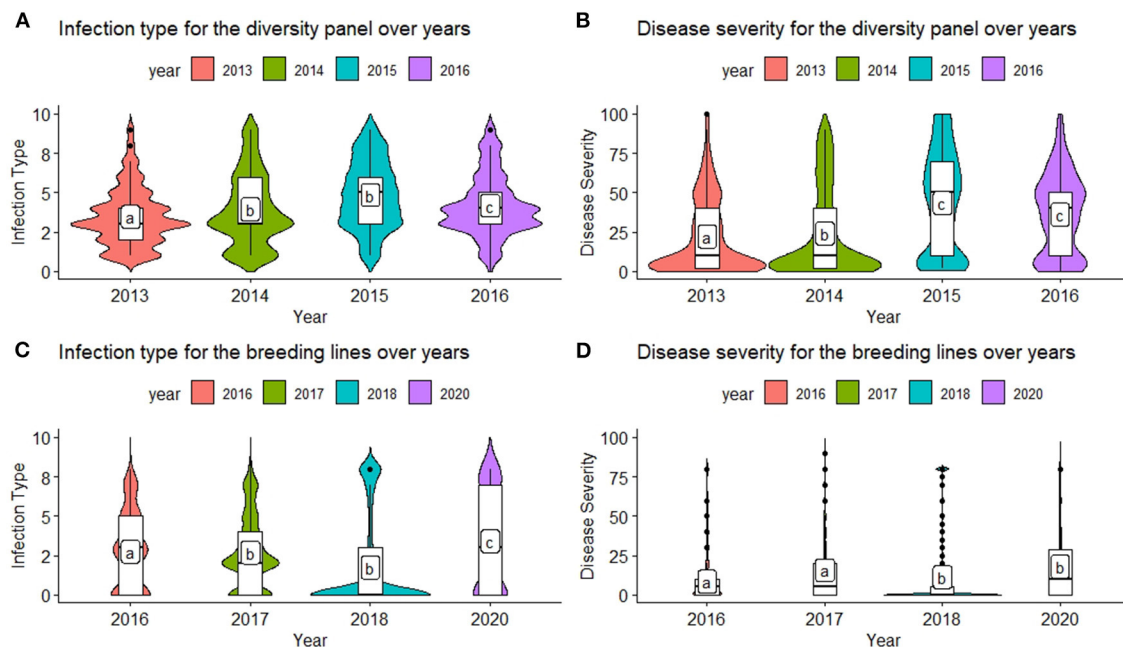


FIGURE 1 | Comparison of infection type (IT) and disease severity (SEV) over years in the diversity panel (DP) lines and breeding lines (BL) training populations using least significant differences. Models labeled with the same letter are not significantly different ($p = 0.05$). **(A)** Infection type for the diversity panel over years. **(B)** Disease severity for the diversity panel over years. **(C)** Infection type for the breeding lines over years. **(D)** Disease severity for the breeding lines over years.

the BL, with SEV having a much higher RMSE for each model than IT (**Supplementary Tables 2, 3**). Further, within SEV for MAS, *Lr68* had a higher RMSE compared to the other markers. *Yr17* did not display a lower RMSE than the other markers, with *Yr10* and ALL_M displaying the lowest RMSE in MAS (**Supplementary Figure 5; Supplementary Tables 4, 5**).

De novo GWAS Markers

The *de novo* GWAS markers increased the prediction accuracy in individual years and across years in the BL, but not in the DP. Only the GWAS_B, GWAS_5, and GWAS_10 sets increased the accuracy with GWAS_25, GWAS_50, and GWAS_100 decreasing the prediction accuracy (**Figures 6, 7**). The largest increase in IT was for GWAS_5 in 2018 with an increase of 0.02 for both IT and SEV (**Supplementary Tables 1, 2; Supplementary Figure 6**). Across years, GWAS_10 had the largest increase of 0.02 in 2016–2018 for SEV (**Supplementary Table 3; Supplementary Figure 7**). The MAS for the *de novo* GWAS markers had larger decreases in MAS compared to the major markers in both the DP and BL (**Supplementary Tables 4, 5**). The larger GWAS sets (GWAS_25, GWAS_50, and GWAS_100) consistently had lower prediction accuracies than the other GWAS sets and the major rust gene markers. GWAS_B using significant markers showed the similar accuracies to GWAS_5, displaying no advantage compared to arbitrarily including markers based on the value of p . The GWAS-GS models displayed a higher RMSE for both GS and MAS in both population and traits compared to the major markers (**Supplementary Figures 8, 9**).

The GWAS_100 sets displayed the highest RMSE out of all models in the cross-validation scenarios with an RMSE of 43.02 (**Supplementary Table 5**).

Validation Sets Major Markers

The validation sets were conducted by combining both training populations and years and predicting the following year as a forward prediction. In doing so, the validation sets were evaluated to demonstrate real-world breeding scenarios wherein all available information was used to create predictions. The first 3 years, 2013–2015, consisted exclusively of the DP, and from 2016 forward, the BL was included due to the availability of training populations. The validation sets resulted in the highest accuracy of all prediction scenarios using the rrBLUP base model, and all major markers reached an accuracy of 0.72 in the SEV for predicting 2014 using the 2013 data (**Figures 8, 9; Supplementary Tables 6, 7**). For the same year, the major markers with the exception of *Yr10* resulted in an increase of the accuracy by 0.01. The major rust markers either performed the same or increased the accuracy for the majority of validation GS predictions.

As a number of environments and years were added to the population, the general prediction accuracy decreased presumably due to the prediction of multiple environments within a year and the inclusion of different training populations. However, as the accuracy decreased for the base rrBLUP model, the effect of fixed markers increased. The largest increase in both cross-validations and validation sets occurred using

TABLE 4 | Frequency of rust-resistant genotypes^a in both the breeding line (BL) and diversity panel (DP) line populations.

| Population | Marker (gene) | Genotype | Numbe ^b | Frequency | Major race effectiveness |
|------------|----------------------------------|----------|--------------------|-----------|--|
| DP | KASP (<i>Yr17</i>) | 0 | 356 | 0.75 | PSTv-322 PSTv-48 PSTv-79 PSTv-11 |
| | | 1 | 30 | 0.06 | |
| | | 2 | 89 | 0.19 | |
| DP | IWB12603 (<i>Qyr.wpg-1B.1</i>) | 0 | 288 | 0.61 | NA ^c |
| | | 1 | 16 | 0.03 | |
| | | 2 | 171 | 0.36 | |
| DP | KASP (<i>Lr68</i>) | 0 | 182 | 0.38 | NA |
| | | 1 | 55 | 0.12 | |
| | | 2 | 238 | 0.50 | |
| DP | Xpsp3000 (<i>Yr10</i>) | 0 | 220 | 0.46 | PSTv-37 PSTv-52 PSTv-322 PSTv-48 PSTv-79 PSTv-11 PSTv-73 |
| | | 1 | 4 | 0.01 | |
| | | 2 | 251 | 0.53 | |
| | | 0 | 1,491 | 0.57 | PSTv-322 PSTv-48 PSTv-79 PSTv-11 |
| | | 1 | 131 | 0.05 | |
| | | 2 | 1,008 | 0.38 | |
| BL | IWB12603 (<i>Qyr.wpg-1B.1</i>) | 0 | 2,244 | 0.85 | NA |
| | | 1 | 53 | 0.02 | |
| | | 2 | 333 | 0.13 | |
| BL | KASP (<i>Lr68</i>) | 0 | 1,255 | 0.48 | NA |
| | | 1 | 166 | 0.06 | |
| | | 2 | 1,209 | 0.46 | |
| BL | Xpsp3000 (<i>Yr10</i>) | 0 | 2,172 | 0.83 | PSTv-37 PSTv-52 PSTv-322 PSTv-48 PSTv-79 PSTv-11 PSTv-73 |
| | | 1 | 11 | 0.00 | |
| | | 2 | 447 | 0.17 | |

^aGenotype: Allele 0: homozygous wild-type allele; Allele 1: heterozygous with both alleles present; Allele 2: homozygous resistant allele.

^bNumber, number of lines.

^cNA, Data were not available.

2013–2017 to predict 2018, resulting in MAS models using *Yr17* and all markers with an increase of 0.17 and 0.16, respectively, for IT and 0.13 and 0.10, respectively, for SEV (Supplementary Tables 8, 9; Supplementary Figure 10). The validation sets were the only prediction scenarios in which MAS performed better than GS models. However, this was not the case for all MAS models, with most major markers showing similar decreases in the accuracy compared to cross-validations.

Additionally, the RMSE was similar to cross-validations with low values for GS models compared to MAS across all markers (Supplementary Figure 11). Further, for SEV, an RMSE for MAS decreased with the addition of years.

De novo GWAS Markers

The *de novo* GWAS marker sets also increased the accuracy when more environments were included. The increase in the

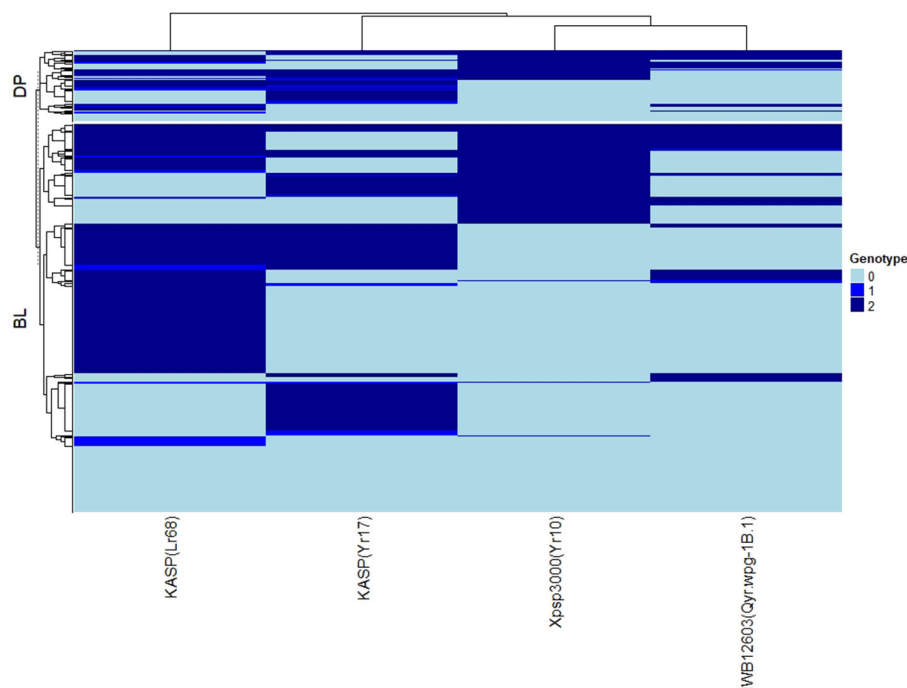


FIGURE 2 | Heat map and hierarchical clustering for lines in the diversity panel (DP) lines and breeding lines (BL) populations for the major rust markers: IWB12603(Qyr.wpg-1B.1), KASP(Lr68), Xpsp3000(Yr10), and KASP(Yr17). Genotype: 0, homozygous wild-type allele; 1, heterozygous with both alleles present; 2, homozygous resistant allele.

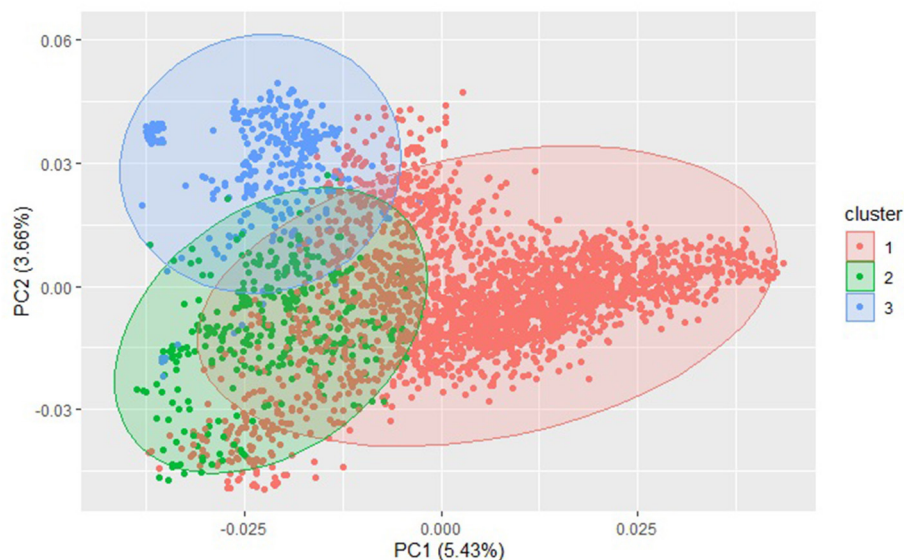
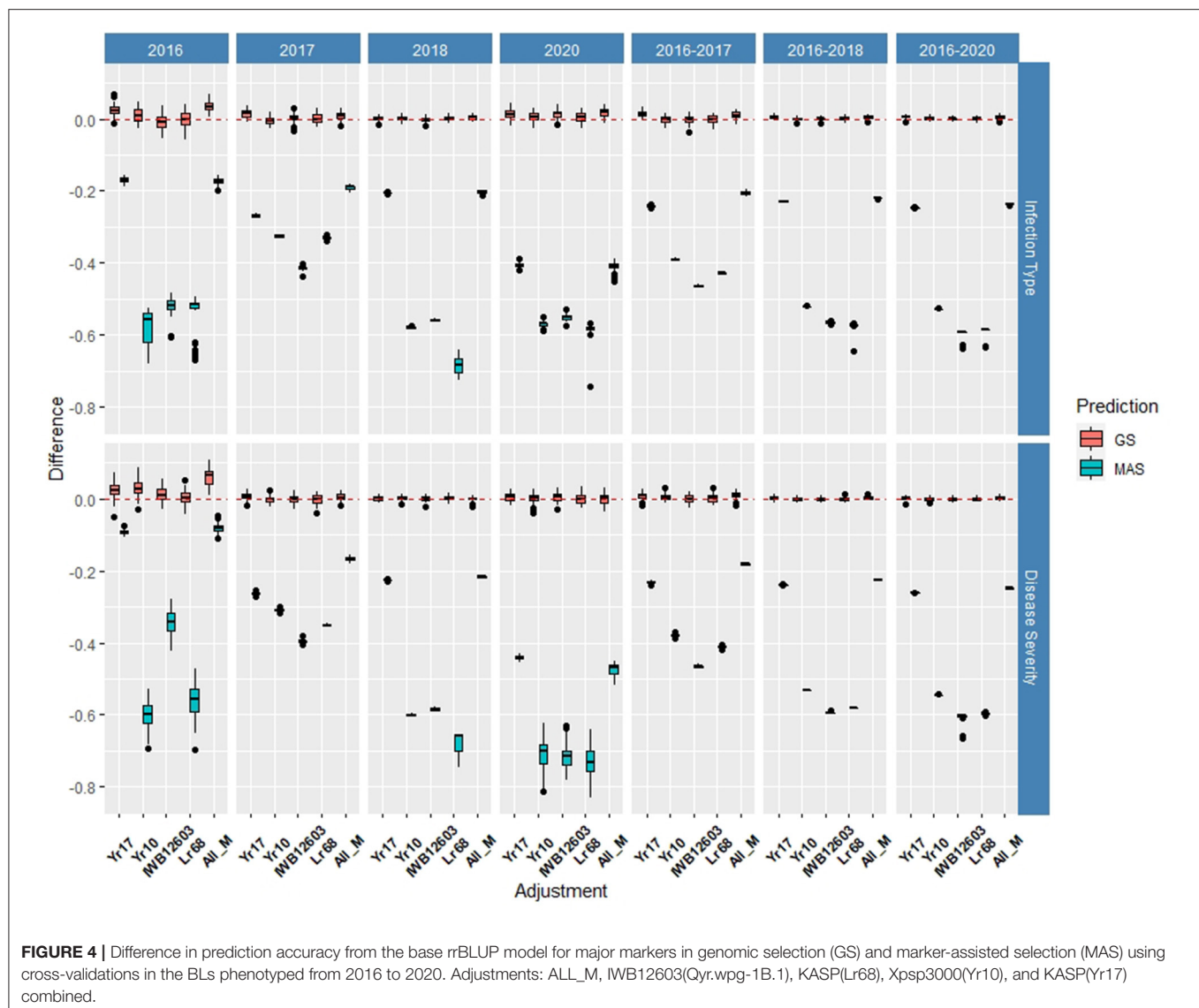


FIGURE 3 | Principal component (PC) biplot and *k*-means clustering of single-nucleotide polymorphism (SNP) genotyped-by-sequencing (GBS) markers from the diversity panel (DP) lines and breeding lines (BL) training populations.

prediction accuracy was not seen in the previous validation sets as seen for the molecular markers for major rust genes. The *de novo* GWAS markers had the largest prediction accuracies in the last two validation sets with GWAS_5 having an accuracy of 0.33 for IT and GWAS having an accuracy of 0.38 for SEV

using 2013–2017 to predict 2018 (**Supplementary Tables 6, 7**). In the last validation set, GWAS had the largest prediction accuracy of 0.55 for IT. Similarly, the smaller GWAS sets had the highest prediction accuracy. In contrast to the cross-validations, the larger GWAS sets did not have a drastic decrease with



GWAS_100, and actually had the same prediction accuracy as the base rrBLUP for IT and an increase of 0.01 for SEV in using 2013–2018 to predict 2020 (**Supplementary Tables 6, 7**). The *de novo* GWAS marker sets had the largest increases in overall scenarios with GWAS_5 having an increase of 0.19 with MAS for IT (**Supplementary Tables 8, 9; Supplementary Figure 12**). Further, MAS for GWAS_100 displayed a much higher RMSE with the highest value for all scenarios reaching an RMSE of 381.71 using 2013–2015 to predict 2016 (**Supplementary Figure 13; Supplementary Table 9**). This prediction scenario was the only scenario using only the DP lines to predict BLs. However, all of the other GS-GWAS sets had an RMSE similar to the major markers for GS and MAS.

Overall Differences

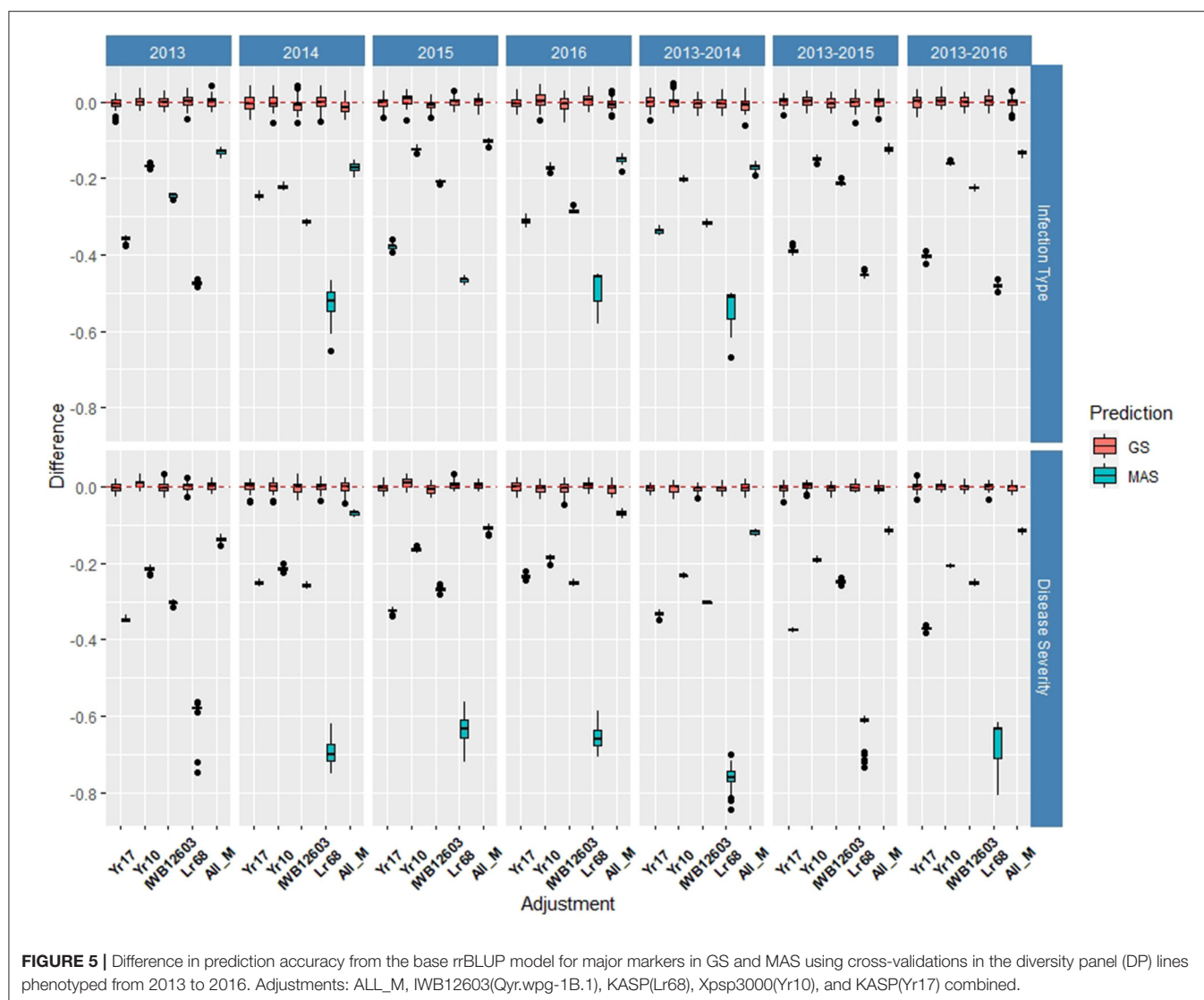
When comparing the different models over all years within each population, we found that the marker for *Yr17* and the

combination of all markers had the largest prediction accuracies. However, the increase was only statistically significant in the BL population and in the validation sets. There was no statistical increase in the prediction accuracy in the DP. The largest mean accuracy in any population was the major rust markers and base rrBLUP for SEV in the DP with an accuracy of 0.64 across all years (**Table 5**). There was also a statistical increase in the prediction accuracy as we increased the combination of years over both IT and SEV training populations with the accuracies of 0.57 and 0.63 for IT and SEV, respectively when years 1–4 were combined (**Table 6**).

DISCUSSION

GS for Disease Resistance

The development of resistant cultivars is the most effective and economical method for controlling diseases such as stripe rust



(Chen and Kang, 2017). Due to the challenges of breeding for both quantitative and qualitative disease resistance, it is recommended to combine them. In addition to the challenges for breeding both major gene qualitative disease resistance and minor gene quantitative resistance are also the common challenges of implementing and integrating any major gene or QTL into new cultivars. These difficulties include inconsistent effects of the QTL due to inconsistent QTL segregations in mapping populations, QTL interaction with genetic background, and QTL interaction with the environment (Bernardo, 2008). However, in addition to the common challenges, qualitative resistance also faces the disadvantage of new virulent races of a pathogen that can overcome major gene resistance (Chen and Kang, 2017). Breeding for minor gene quantitative resistance tends to produce a more durable resistance in BLs because it relies on multiple small-effect alleles. Similar to other agronomic traits, breeding for quantitative resistance requires multiple breeding cycles to gradually improve resistance (Poland and

Rutkoski, 2016). The lack of qualitative resistance durability coupled with the challenge in identifying and breeding for quantitative resistance creates a unique opportunity for GS to identify quantitative resistance by accounting for minor-effect genes in the presence of large-effect major genes.

The goal of this study was to identify the best GS method for disease resistance in the presence of both major and minor genes. In our study, we used stripe rust as an example of the disease with both major and minor resistant genes. Previous studies on the GS of stripe rust showed promising prediction accuracies. Muleta et al. (2017) showed that accuracy increased with population size and marker density and reached up to 0.80. Ornella et al. (2012) reported accuracies in The International Maize and Wheat Improvement Center (CIMMYT) wheat populations of values greater than 0.50 for stripe rust, but showed a lower accuracy when compared to stem rust. In our study, the prediction accuracy for both IT and SEV reached an accuracy of up to 0.67 and 0.69 in cross-validations, respectively. Further, IT and

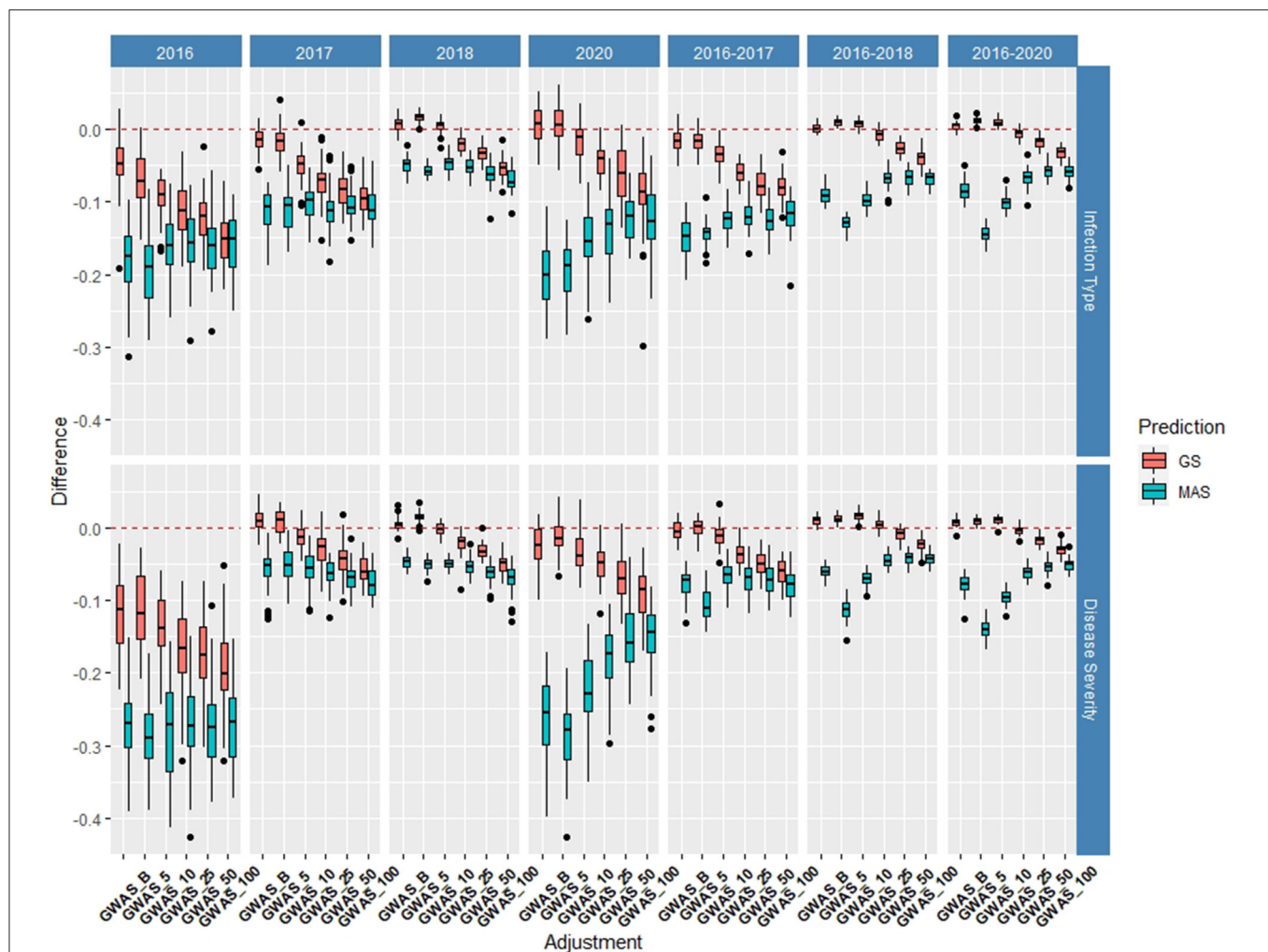


FIGURE 6 | Difference in prediction accuracy from the base rrBLUP model for *de novo* genome-wide association study (GWAS) markers in GS and MAS using cross-validations in the BLs phenotyped from 2016 to 2020. Adjustments: GWAS_B, genome-wide association study assisted GS (GWAS-GS) with Bonferonni significant markers; GWAS_5, GWAS-GS with the top five significant markers; GWAS_10, GWAS-GS with the top 10 significant markers; GWAS_25, GWAS-GS with the top 25 significant markers; GWAS_50, GWAS-GS with the top 50 significant markers; GWAS_100, GWAS-GS with the top 100 significant markers.

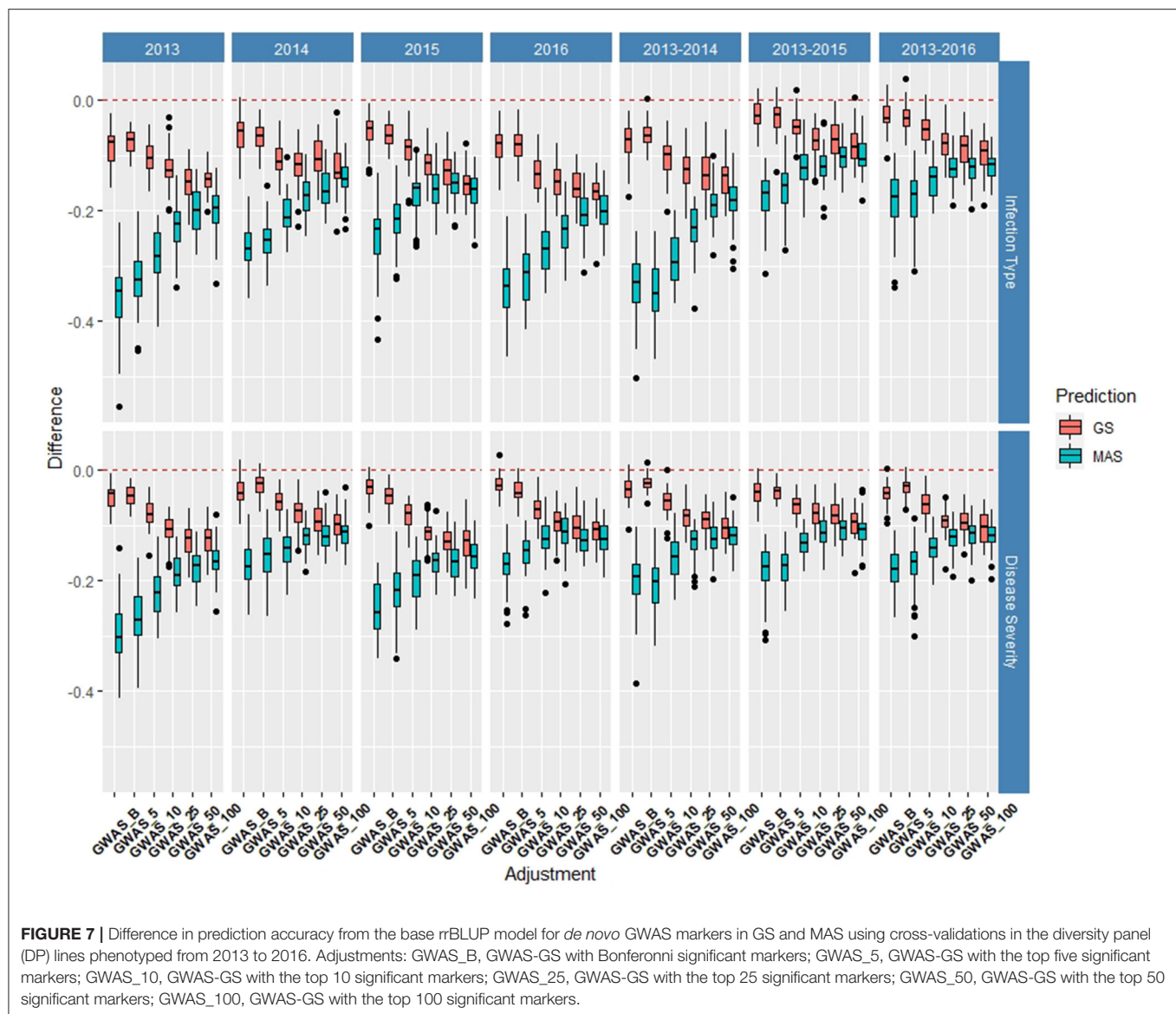
SEV reached the accuracies of up to 0.66 and 0.72 in validation sets, respectively. In comparison to other rust diseases, Rutkoski et al. (2014, 2015) showed promising results to predict stem rust with the accuracies of up to 0.50. Overall, our study showed high prediction accuracies in comparison to most rust prediction studies, and further displayed the feasibility for accurately predicting disease resistance in the presence of major and minor resistant genes.

Major Markers

When major genes are present, a large portion of the genetic variance for a trait may be due to the unknown QTL with minor effects (Bernardo, 2014). The other minor-effect QTL will not necessarily be integrated when major genes are integrated into cultivars. The lack of integration can be attributed to not being able to use MAS and the difficulties outlined previously in

pyramiding major-effect genes. In contrast, GS simultaneously models all QTL (Meuwissen et al., 2001). However, the use of GS models such as rrBLUP will underestimate the effect of the major QTL. Therefore, the inclusion of the major-effect QTL as fixed effects can increase accuracy. According to Bernardo (2014), major genes should be used in prediction models when only a few major genes are present, and each gene accounts for more than 10% of the variation.

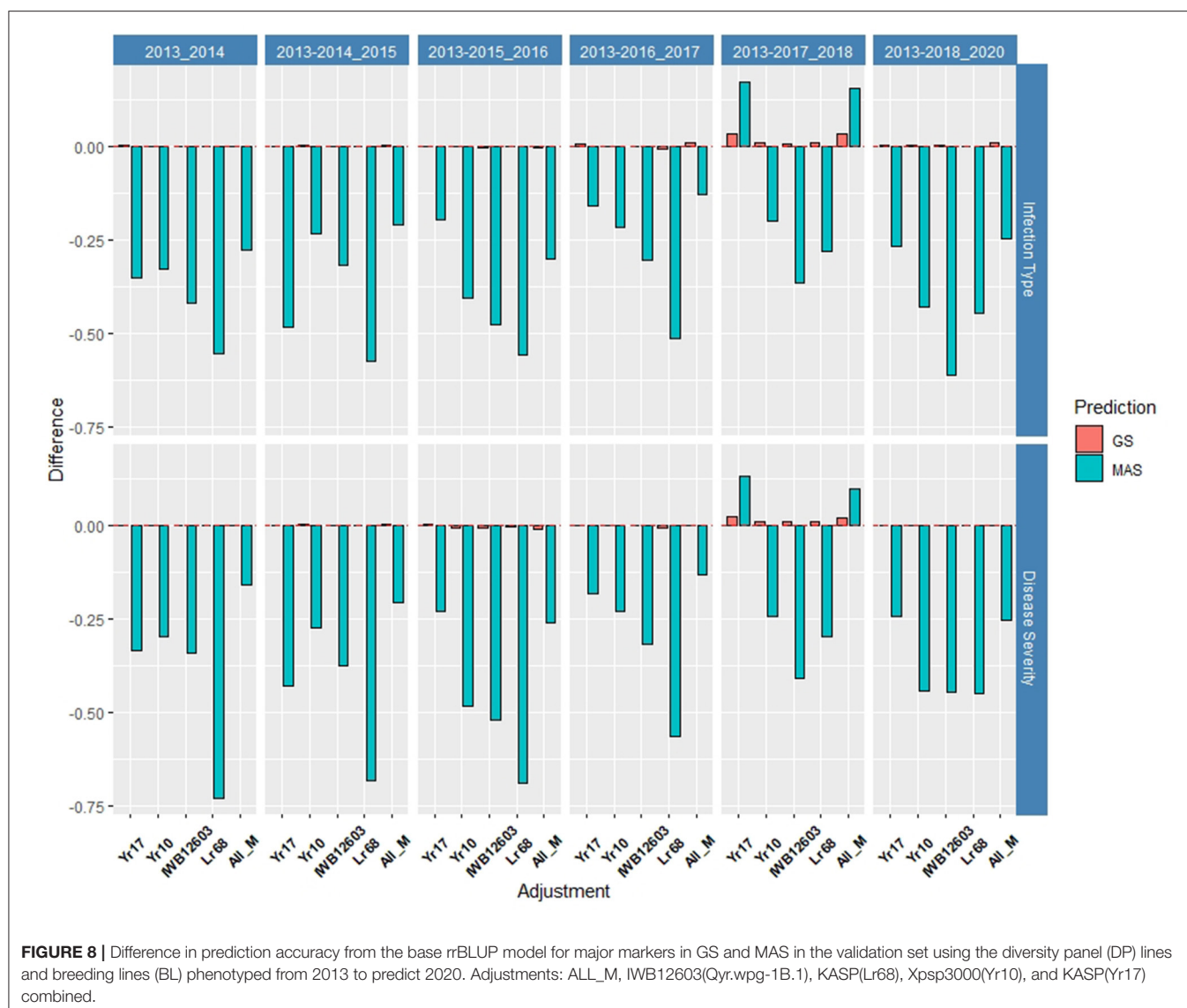
In this study, the major gene in both populations was *Yr17*. In the BL, *Yr17* accounted for up to 0.40 prediction accuracy when used in MAS, and therefore accounts for a large amount of variation. The moderate accuracy of *Yr17* supports that even with the degradation of the ASR for *Yr17*, it still provides resistance for APR as indicated in Liu et al. (2018). The other major rust genes present in the BL would be considered as minor-effect genes with a near-zero prediction accuracy within MAS or, in the



case of the marker for *Yr10*, only produced an accuracy greater than 0.10 in a few prediction scenarios. The higher accuracy in the BL for *Yr17* also shows a lower RMSE compared to the other markers. However, within the DP, all of the markers with the exception of *Lr68* produced accuracies greater than 0.20, with *IWB12603* reaching 0.34 and *Yr10* reaching the highest accuracies for MAS within cross-validations of 0.42, and could be considered as major-effect markers. Additionally, the higher accuracy for *Yr10* was coupled with a lower RMSE than the other markers.

Even with the moderate accuracies of the major rust markers in MAS, we observed only a slight increase in the prediction accuracy when the major markers were included in our GS models, and relatively a lower RMSE than the MAS. The major markers only increased the prediction accuracy at a maximum of 0.06 within the cross-validation scenarios and 0.03 within

the validation sets. Interestingly, the validation sets resulted in the highest accuracy of all scenarios with 0.72 for the base GS model and the inclusion of the major markers when predicting SEV in 2014 using 2013 with the small RMSE values for all markers. These results are in direct contrast to previous studies showing a higher accuracy in cross-validations (Lozada and Carter, 2019; Merrick and Carter, 2021). Validation sets are a more realistic approach for GS because it is comparable to how GS would be implemented in breeding programs (Lozada and Carter, 2019). However, the major markers only increased the prediction accuracy as the overall prediction accuracy decreased. For example, using 2013–2017 to predict 2018, all of the major markers increased the prediction accuracy, but the base prediction was only 0.27, and the markers increased the accuracy by 0.03 maximum in all scenarios. Further, the major markers had much larger increases in the MAS scenarios with a maximum



increase of 0.17, but resulted in a higher RMSE. Therefore, the inclusion of the major markers provides an advantage in the more realistic validation sets when the base GS model has poor predictive ability.

In the context of GS models and breeding programs, a small increase in the prediction accuracy would be considered negligible in realistic breeding scenarios. The results in our study are in contrast to previous studies showing that the major markers had a large increase in prediction accuracies in GS models for other diseases such as stem rust (Rutkoski et al., 2014) and Fusarium head blight (Arruda et al., 2016). One hypothesis for the lack of increase in the prediction accuracy may be due to GS models accounting for a majority of variation in both the major- and minor-effect markers for disease resistance, or the major-effect markers may be accounted for in the models of the GBS markers. However, the ridge regression penalty reduces the effect of large-effect markers, hence the additional

variation would need to be accounted for by other small-effect markers (Rice and Lipka, 2019). Additionally, the lack of increase in the prediction accuracy may be due to the major markers not accounting for enough phenotypic variation. Due to the reduction in the effect of the major markers, Bernardo (2014) suggested implementing markers that account for more than 10% of the variation as mentioned previously. This theory may be disproved by the major markers that display a moderate accuracy in MAS models. However, this may be the case for *Lr68*, which displayed a minimal effect in both MAS models and GS models.

Further, the lack of increase in the prediction accuracy may be beneficial in demonstrating that other uncharacterized resistant QTL can still provide a large amount of disease resistance within the populations either alone or in conjunction with major genes. In this case, our results would be beneficial in confirming the presence of minor-effect QTL for quantitative resistance and provide a more durable resistance within the

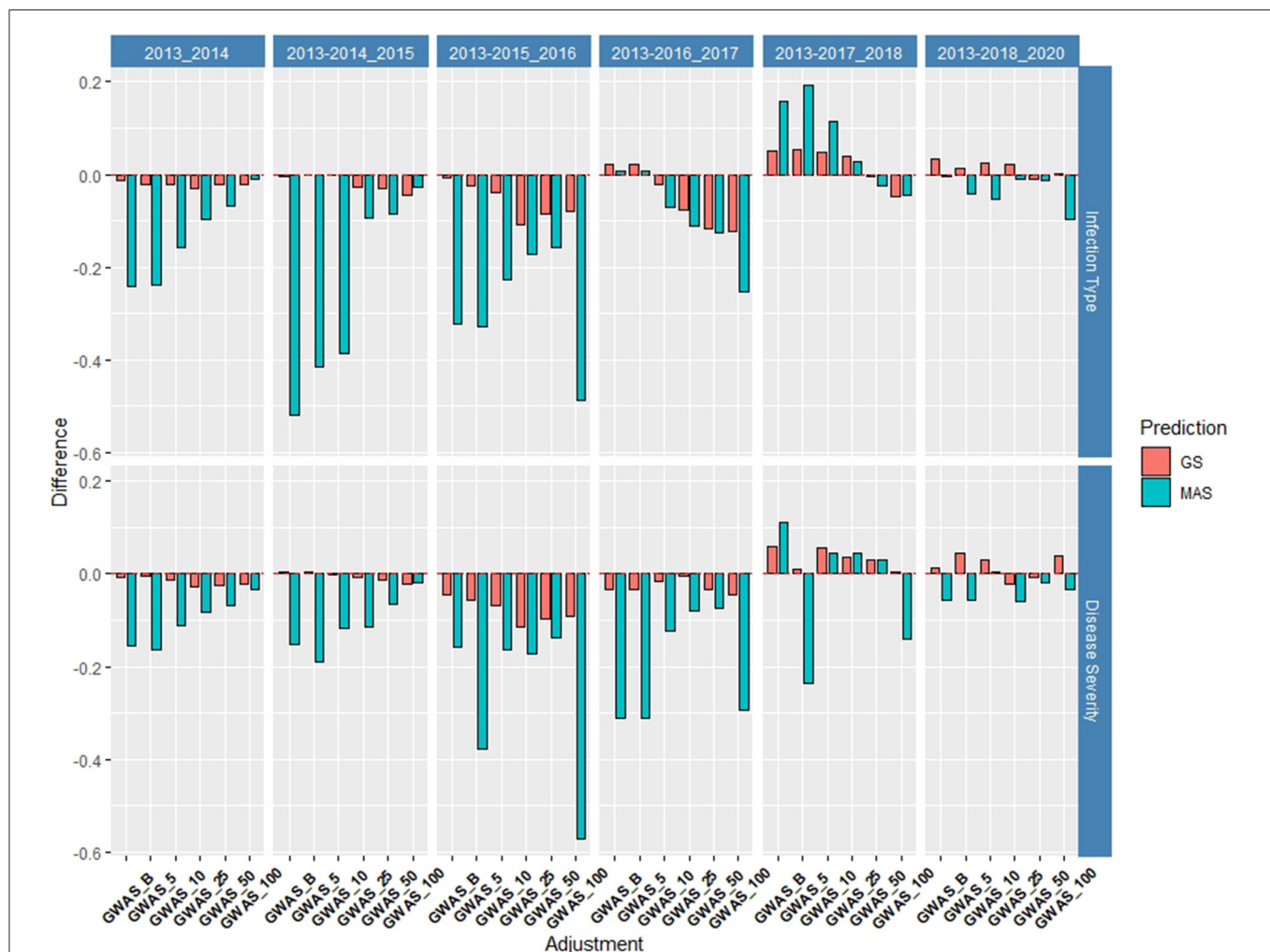


FIGURE 9 | Difference in prediction accuracy from the base rrBLUP model for *de novo* GWAS markers in GS and MAS in the validation set using the diversity panel (DP) lines and breeding lines (BL) phenotyped from 2013 to predict 2020. Adjustments: GWAS_B, GWAS-GS with Bonferonni significant markers; GWAS_5, GWAS-GS with the top five significant markers; GWAS_10, GWAS-GS with the top 10 significant markers; GWAS_25, GWAS-GS with the top 25 significant markers; GWAS_50, GWAS-GS with the top 50 significant markers; GWAS_100, GWAS-GS with the top 100 significant markers.

training populations. Therefore, we can conclude that genotyping and selecting major genes for disease resistance may not be necessary when the breeding programs can use more cost-effective genome-wide markers to implement GS with more consistent results.

De novo GWAS Markers

Frequently, the major markers for disease resistance are either unknown or have an uncharacterized effect within the populations. Therefore, GWAS can be performed to characterize disease-resistant QTL within a population, and the significant markers can be used as fixed-effect covariates (Rice and Lipka, 2019). In Zhang et al. (2014), publicly available GWAS markers were integrated into prediction models but only increased the accuracy by 0.01, similar to our results. In contrast, we used *de novo* GWAS markers dependent on the training population. This

approach has been used for FHB in which Arruda et al. (2016) demonstrated an increase in the accuracy of up to 0.14. These results were also demonstrated in Spindel et al. (2016), in which *de novo* GWAS markers implemented into GS increased the accuracies more than 0.10 in rice (*Oryza sativa* L.). However, in our study, the *de novo* GWAS markers only marginally increased the accuracy, or in the case of implementing more than 25 markers, decreased accuracy in the majority of cross-validation scenarios and an increased RMSE. A reduction in the prediction accuracy and an increase of RMSE with a larger set of *de novo* GWAS markers may be attributed to an increase in the bias of the model and an increase of RMSE due to overfitting as seen in Raymond et al. (2018) or due to the difficulty experienced by the model to simultaneously estimate all of the fixed effects (Bernardo, 2014). A reduction in the prediction accuracy was also shown in Rice and Lipka (2019).

TABLE 5 | Comparison of genomic selection models accuracy and pairwise comparisons for stripe rust IT and disease severity (SEV) for PNW winter wheat DP lines and BLs phenotyped from 2013 to 2020 in Central Ferry, Lind, and Pullman, WA, USA over all individual population cross-validation sets and combined validation sets.

| Trait | Pop | rrBLUP | Yr17 | Yr10 | IWB12603 | Lr68 | All_M | GWAS_B | GWAS_5 | GWAS_10 | GWAS_25 | GWAS_50 | GWAS_100 |
|-------|-----|--------|--------|---------|----------|---------|--------|--------|--------|---------|---------|---------|----------|
| IT | BL | 0.58b | 0.59a | 0.58b | 0.58b | 0.58b | 0.59a | 0.57c | 0.57c | 0.56d | 0.54e | 0.52f | 0.51g |
| | DP | 0.53ab | 0.53ab | 0.53a | 0.52b | 0.53ab | 0.52b | 0.47c | 0.47c | 0.44d | 0.42e | 0.41f | 0.4g |
| | VS | 0.48d | 0.49b | 0.49cd | 0.48d | 0.48d | 0.49ab | 0.5a | 0.49bc | 0.48d | 0.45e | 0.44f | 0.43g |
| SEV | BL | 0.59b | 0.6ab | 0.59ab | 0.59b | 0.59b | 0.6a | 0.57cd | 0.58c | 0.56d | 0.55e | 0.53f | 0.52g |
| | DP | 0.64ab | 0.64ab | 0.64a | 0.63b | 0.64ab | 0.63ab | 0.6c | 0.6c | 0.57d | 0.54e | 0.54f | 0.53g |
| | VS | 0.54bc | 0.55a | 0.54abc | 0.54abc | 0.54abc | 0.55ab | 0.54cd | 0.54d | 0.54cd | 0.52e | 0.52e | 0.52e |

Models labeled with the same letter are not significantly different ($p = 0.05$). Adjustments: ALL_M: IWB12603(Qrwpwg-1B.1), KASP (Lr68), Xpsp3000(Yr10), and KASP (Yr17) combined; GWAS_B, genome-wide association assisted genomic selection (GWAS-GS) with Bonferroni significant markers; GWAS_5, GWAS-GS with the top 5 significant markers; GWAS_10, GWAS-GS with the top 10 significant markers; GWAS_25, GWAS-GS with the top 25 significant markers; GWAS_50, GWAS-GS with the top 50 significant markers; GWAS_100, GWAS-GS with the top 100 significant markers.

TABLE 6 | Comparison of the number of years in the training populations on overall GS model accuracy and pairwise comparisons for stripe rust infection type (IT) and disease severity (SEV) for PNW winter wheat over both the DP lines and BLs phenotyped from 2013 to 2020 in Central Ferry, Lind, and Pullman, WA, USA in the cross-validation sets.

| Trait | Year 1 | Year 2 | Year 3 | Year 4 | Year 1–2 | Year 1–3 | Year 1–4 |
|-------|--------|--------|--------|--------|----------|----------|----------|
| IT | 0.52d | 0.43f | 0.57a | 0.54c | 0.49e | 0.56b | 0.57a |
| SEV | 0.53f | 0.55e | 0.61b | 0.59c | 0.58d | 0.62a | 0.63a |

Models labeled with the same letter are not significantly different (P -value = 0.05).

Another hypothesis may be stated for why the *de novo* GWAS markers failed to increase the prediction accuracy due to the inclusion of false positives within GWAS models. To mitigate this, we included a GWAS-GS model that only included significant markers based on a Bonferroni correction of 0.05. However, this model failed to self-differentiate from another smaller set of GWAS-GS models. The lack of reduction was mainly seen in our cross-validation sets. Within cross-validation, the training population is divided. The division of the training populations may be one cause of the lack of increase of the prediction accuracy. The smaller validation fold within a cross-validation may have a weak association with the markers found in the larger training folds, as hinted at by Rice and Lipka (2019). The weak association theory may be supported by the contrasting results seen in the validation sets.

Similar to the inclusion of the major markers in the cross-validations, the validation sets showed an increase in the prediction accuracy when the *de novo* GWAS markers were included and displayed the largest increases from GS models. The GWAS model with significant markers only (GWAS_B) displayed the largest increase of 0.06 in the SEV. Once again, this increased prediction accuracy was observed as the prediction accuracy of the base GS model decreased. This occurrence in both the major and *de novo* GWAS markers demonstrates the ability to increase and maintain a high accuracy as the GS model fails in predicting lines. Therefore, we can conclude that even though fixed-effect markers may not increase the accuracy in typical cross-validation scenarios, they are beneficial in more realistic validation set approaches similar to the major markers.

However, similar to the major markers, increased prediction accuracy with the inclusion of *de novo* GWAS markers was very small relative to the high accuracy for most scenarios. Further, small sets of *de novo* GWAS markers were similar in consistency to the major markers. Therefore, there is little benefit in characterizing major-effect disease resistance markers for GS over implementing the GWAS-GS methods that would use the same sets of markers like GS models.

Training Population and Environment

We compared the effect of the major and *de novo* GWAS markers in different training populations that are commonly used in breeding programs. The frequency and source of both major disease- and minor disease-resistant genes vary. For instance, the BL population consists of WSU BLs that have been selected for resistance, specifically for *P. striiformis* f. sp. *tritici* races in Washington, and therefore has a high level of resistance

throughout the population. In comparison, the DP consists of varieties from various breeding programs in the PNW. The sources of resistance in the varieties are more similar within the BL than in the DP, with the DP containing major genes different from the major markers chosen in this study common in the WSU germplasm or selected for resistance to races not present in eastern Washington.

The differences in the frequency of major genes were observed in the major rust markers used in this study. In the BL, the *Yr17* marker showed an increase in the prediction accuracy for GS models and a relatively high accuracy in MAS models compared to the other markers. However, this was not consistently seen in the DP. The inconsistent effect of *Yr17* in different training populations may be due to the higher frequency of *Yr17* in the BL compared to the DP. This may also be supported by the higher accuracies for *Yr10* and *IWB12603* in the DP compared to the BL, and both of these rust genes have a higher frequency in the DP than in the BL. Our study showed that regardless of the frequency of the rust-resistant genotypes, there was only a small to nil increase in the prediction accuracy. Therefore, GS would be more accurate than MAS regardless of the frequency of the known rust-resistant genotypes in a breeding program due to the ability to account for both major disease and minor disease-resistant genes.

In addition to different frequencies of major genes, the general composition of the training populations can affect GS prediction accuracy (Asoro et al., 2011). The composition of the training population affects the accuracy due to both population structure and genetic relatedness (Habier et al., 2007; Asoro et al., 2011; Mirdita et al., 2015). We compared the population structure in our models by plotting principal components and identified three clusters indicating distinct subpopulations. In addition, the population structure was not taken into account in our GS models. However, we can see the effect of genetic relatedness and population in both our cross-validation and validation sets. The BL had a statistically higher mean accuracy for both IT and SEV than the DP in cross-validations, which could be attributed to the closer genetic relatedness of the population and sources of resistance as mentioned previously. A higher prediction accuracy for the BL is advantageous for breeding programs because they can use their existing breeding trials for GS without screening a DP outside their breeding program. In the validation sets, we see an initial increase in the accuracy due to the DP being the only population in the training populations, but as we added in BLs, the accuracy decreased. The accuracy was reduced when the DP predicted the BL, but eventually increased as more BLs were introduced into the training population. The decrease in validation sets can also be attributed to GEI (Michel et al., 2016; Huang et al., 2018; Lozada and Carter, 2019, 2020; Haile et al., 2020).

Further, GEI is important for qualitative disease resistance. Race-specific qualitative resistance is dependent on the race in the environment and thus can lead to larger environmental effects (Poland and Rutkoski, 2016). In contrast, GEI has a much smaller effect on minor-gene quantitative resistance due to the lack of a gene-for-gene interaction. In our study, the most frequent races were similar from year to year, and therefore may not be a significant factor in the differing prediction accuracy.

In this study, disease resistance screening was dependent on the natural occurrence of stripe rust for disease pressure, and therefore the overall effect of the environment is important. Additionally, diseases such as stripe rust are affected by several environmental factors, including moisture, temperature, and wind. Further, disease SEV is affected by the other aspects of the disease triangle, disease inoculum, and a susceptible host to induce disease development (Chen, 2005). Disease development and the quality of the phenotypic data obtained from the unreplicated trials may also explain the differences in the prediction accuracy from year to year, especially in the DP in which the same lines are phenotyped every year. Meanwhile, BLs are only phenotyped in a single year, and therefore the difference from 1 year to the next can be either disease incidence as in the DP or the changes occurred due to differing levels of resistance within BLs. In addition, we see an increase in the prediction accuracy for both the cross-validation and validation sets as we increase the number of environments within our training population. The increase in accuracy may be accounted for by the inclusion of GEI within our phenotypic adjustments and GS models as reported in previous studies, as well as the general high heritability for disease resistance (Crossa et al., 2014; Jarquín et al., 2014; Haile et al., 2020; Merrick and Carter, 2021). Overall, our GS models accurately predicted disease resistance in different training populations and environments, and therefore will be an important strategy for selecting for disease resistance.

Applications in Breeding

Genome selection is beneficial for complex traits and can outperform phenotypic selection and MAS for low heritable traits. However, there may be little benefit in using GS for selection purposes for highly heritable traits such as disease resistance (Poland and Rutkoski, 2016). In the case of highly heritable traits, GS can still outperform phenotypic selection and MAS in terms of gain per unit time when implemented in the early stages of the breeding cycle (Bernardo and Yu, 2007; Rutkoski et al., 2011). In our study, a high prediction accuracy would allow an increase in genetic gain by decreasing the cycle time of the breeding program and rapidly accumulating favorable alleles for disease resistance (Rutkoski et al., 2011).

Even though phenotypic selection has been successfully implemented for disease resistance, without controlled experiments, one cannot determine whether the resistance is quantitative or qualitative. Therefore, we cannot conclude whether the resistance will be durable in the long term. Alternatively, we can implement MAS to select qualitative and quantitative disease resistance within the BLs to bypass the need for controlled experiments. However, as seen in our study, MAS does not account for all of the resistance within the lines in either of the training populations, as shown by a decrease in the prediction accuracy for MAS models. MAS also has limitations when it comes to pyramiding multiple markers, as discussed previously, and is a form of tandem selection (Bernardo, 2014). In contrast, GS is a form of selection index and has been shown to be superior to tandem selection (Hazel and Lush, 1942). Using GS, we can select for the accumulation of all-resistant QTL to take advantage of the quantitative and qualitative resistant genes within a population,

even when they are uncharacterized. Furthermore, by using fixed effects, we can select the lines that have a major marker of interest (Poland and Rutkoski, 2016). Therefore, GS will have a place in selecting for both quantitative and qualitative disease resistance.

Another advantage in implementing GS is by reducing both genotyping and phenotyping within a breeding program. GS can remove the need for genotyping for major and minor genes for selection purposes. This is further supported by the similar accuracies between major and *de novo* GWAS markers. By utilizing genome-wide markers, we can not only implement GS or GWAS-GS but also utilize the markers for additional traits, thus making the genome-wide markers more cost-effective (Poland and Rutkoski, 2016). Likewise, with the help of GS, breeding programs can reduce the need for phenotypic screening in disease nurseries in multiple locations and free up resources for screening more lines and increase genetic gain (Poland and Rutkoski, 2016).

Furthermore, the challenges introduced by the environment mentioned previously provide another advantage in using GS for disease resistance. GS models will help select cultivars with durable quantitative resistance with the accumulation of favorable alleles and select for disease resistance in environments not conducive to disease incidence needed for phenotypic selection. Overall, the high accuracy of GS models in our study displays the ability to predict durable disease resistance and account for uncharacterized minor-effect QTL in the presence of known major genes.

CONCLUSIONS

This study showed the ability to accurately predict disease resistance using major and minor genes. The small to nil increase in the prediction accuracy for the major markers indicates the need for a careful selection of the major markers that account for a large variation in the training and test populations. Further, a comparison of the number of *de novo* GWAS markers shows that a small number of *de novo* GWAS markers should be used instead of a large set of markers to keep from overfitting the model. Additionally, fixed-effect markers may not provide a benefit in scenarios with already high prediction accuracy. However, in prediction scenarios with low accuracies such as in more realistic validation sets, the inclusion of both major markers and *de novo* GWAS helps to account for a variation in case of the failure of the base GS models. Moreover, we can increase the accuracy with the inclusion of additional environments and by using the populations that are genetically related such as the BL. Overall, there were no disadvantages in the inclusion of the major or

de novo GWAS markers. The lack of increase of the prediction accuracy with the inclusion of fixed effects coupled with a large decrease in the accuracy using MAS indicates the presence of minor-effect QTL for quantitative resistance and thus durable resistance within the training populations. This study showed the ability to predict disease resistance and accumulate favorable alleles for durable disease resistance in the presence of major and minor resistance genes.

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 at: <https://github.com/lfmerrick21/Major-and-Minor-Genes>.

AUTHOR CONTRIBUTIONS

LM conceptualized the idea, analyzed the data, and drafted the manuscript. AB genotyped the KASP markers, reviewed, and edited the manuscript. XC reviewed and edited the manuscript. AC supervised the study, conducted field trials, edited the manuscript, and obtained the funding for the project. All authors contributed to the article and approved the submitted version.

FUNDING

This research was partially funded by the National Institute of Food and Agriculture (NIFA) of the US Department of Agriculture (Award number 2016-68004-24770), Hatch project 1014919, and the O.A. Vogel Research Foundation at WSU.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Washington State University Winter Wheat Breeding Program personnel Gary Shelton and Kyall Hagemeyer for plot maintenance and data collection under field conditions. We would also like to thank Gina Brown-Guedira, Jared Smith, Brian Ward, and staff at the Eastern Regional Small Grains Genotyping Laboratory for their assistance with DNA library prep and GBS sequencing and analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.713667/full#supplementary-material>

REFERENCES

- Agrios, G. N. (2005). *Plant Pathology*. Amsterdam: Elsevier Academic Press.
- 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. doi: 10.1126/science.aar7191
- Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., et al. (2016). Comparing genomic selection and marker-assisted selection for Fusarium head blight resistance in wheat (*Triticum aestivum* L.). *Mol. Breed.* 36:84. doi: 10.1007/s11032-016-0508-5
- Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., and Jannink, J.-L. (2011). Accuracy and training population design for genomic selection

- on quantitative traits in elite North American oats. *Plant Genome* 4:132. doi: 10.3835/plantgenome2011.02.0007
- Bai, B., He, Z. H., Asad, M. A., Lan, C. X., Zhang, Y., Xia, X. C., et al. (2012). Pyramiding adult-plant powdery mildew resistance QTLs in bread wheat. *Crop Pasture Sci.* 63, 606–611. doi: 10.1071/CP12183
- Bariana, H. S., Brown, G. N., Ahmed, N. U., Khatkar, S., Conner, R. L., Wellings, C. R., et al. (2002). Characterisation of *Triticum vavilovii*-derived stripe rust resistance using genetic, cytogenetic and molecular analyses and its marker-assisted selection. *Theor. Appl. Genet.* 104, 315–320. doi: 10.1007/s001220100767
- Bernardo, R. (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci.* 48:1649. doi: 10.2135/cropsci2008.03.0131
- Bernardo, R. (2014). Genomewide selection when major genes are known. *Crop Sci.* 54:68. doi: 10.2135/cropsci2013.05.0315
- Bernardo, R., and Yu, J. (2007). Prospects for genomewide selection for quantitative traits in maize. *Crop Sci.* 47:1082. doi: 10.2135/cropsci2006.11.0690
- Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi: 10.1016/j.ajhg.2018.07.015
- Case, A. J., Naruoka, Y., Chen, X., Garland-Campbell, K. A., Zemetra, R. S., and Carter, A. H. (2014). Mapping stripe rust resistance in a Brundage X Coda winter wheat recombinant inbred line population. *PLoS ONE* 9:e91758. doi: 10.1371/journal.pone.0091758
- Chen, X. (2013). High-temperature adult-plant resistance, key for sustainable control of stripe rust. *Am. J. Plant Sci.* 4, 608–627. doi: 10.4236/ajps.2013.43080
- Chen, X., and Kang, Z. (2017). *Stripe Rust*. Dordrecht: Springer. doi: 10.1007/978-94-024-1111-9
- Chen, X., and Line, R. F. (1995). Gene action in wheat cultivars for durable, high-temperature, adult-plant resistance and interaction with race-specific, seedling resistance to *Puccinia striiformis*. *Phytopathology* 85, 567–572. doi: 10.1094/Phyto-85-567
- Chen, X., Washington, S. U., and Line, R. F. (1995). Gene number and heritability of wheat cultivars with durable, high-temperature, adult-plant (HTAP) resistance and interaction of HTAP and race-specific seedling resistance to *Puccinia striiformis*. *Phytopathology* 85, 573–578. doi: 10.1094/Phyto-85-573
- Chen, X. M. (2005). Epidemiology and control of stripe rust [*Puccinia striiformis* f. sp. tritici] on wheat. *Can. J. Plant Pathol.* 27, 314–337. doi: 10.1080/0706060509507230
- Crossa, J., Pérez, P., Hickey, J., Burguño, J., Ornela, L., Cerón-Rojas, J., et al. (2014). Genomic prediction in CIMMYT maize and wheat breeding programs. *Heredity* 112, 48–60. doi: 10.1038/hdy.2013.16
- Cullis, B. R., Smith, A. B., and Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *J. Agric. Biol. Environ. Stat.* 11:381. doi: 10.1198/108571106X154443
- de Mendiburu, F., and de Mendiburu, M. F. (2019). Package “agricolae.” *R Package Version* 1, 2–8.
- 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
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4:250. doi: 10.3835/plantgenome2011.08.0024
- Federer, W. F. (1956). *Experimental Design, Theory and Application*. New York, NY: Macmillan. doi: 10.1097/00010694-195604000-00015
- Frick, M. M., Huel, R., Nykiforuk, C. L., Conner, R. L., Kuzyk, A., and Laroche, A. (1998). “Molecular characterisation of a wheat stripe rust resistance gene in Moro wheat,” in *Proceedings of the 9th International Wheat Genetics Symposium* (Saskatoon, SK: University Extension Press, University of Saskatchewan Saskatoon), 181–182.
- Gerechter-Amitai, Z. K., Van Silfhout, C. H., Grama, A., and Kleitman, F. (1989). Yr 15—a new gene for resistance to *Puccinia striiformis* in *Triticum dicoccoides* sel. G-25. *Euphytica* 43, 187–190. doi: 10.1007/BF00037912
- Goldman, I. (2019). *Plant Breeding Reviews*. Chichester: John Wiley and Sons. doi: 10.1002/9781119616801
- González-Camacho, J. M., Ornela, L., Pérez-Rodríguez, P., Gianola, D., Dreisigacker, S., and Crossa, J. (2018). Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. *Plant Genome* 11:170104. doi: 10.3835/plantgenome2017.11.0104
- Habier, D., Fernando, R. L., and Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177, 2389–2397. doi: 10.1534/genetics.107.081190
- Haile, T. A., Walkowiak, S., N'Diaye, A., Clarke, J. M., Hucl, P. J., Cuthbert, R. D., et al. (2020). Genomic prediction of agronomic traits in wheat using different models and cross-validation designs. *Theor. Appl. Genet.* 20, 1–18. doi: 10.1007/s00122-020-03703-z
- Hazel, L. N., and Lush, J. L. (1942). The efficiency of three methods of selection. *J. Hered.* 33, 393–399. doi: 10.1093/oxfordjournals.jhered.a105102
- Helguera, M., Khan, I. A., Kolmer, J., Lijavetzky, D., Zhong-Qi, L., and Dubcovsky, J. (2003). PCR assays for the Lr37-Yr17-Sr38 cluster of rust resistance genes and their use to develop isogenic hard red spring wheat lines. *Crop Sci.* 43, 1839–1847. doi: 10.2135/cropsci2003.1839
- Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., Rosewarne, G. M., Periyannan, S. K., Viccars, L., et al. (2012). Lr68: a new gene conferring slow rusting resistance to leaf rust in wheat. *Theor. Appl. Genet.* 124, 1475–1486. doi: 10.1007/s00122-012-1802-1
- Huang, M., Liu, X., Zhou, Y., Summers, R. M., and Zhang, Z. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience* 8:giy154. doi: 10.1093/gigascience/giy154
- Huang, M., Ward, B., Griffey, C., Van Sanford, D., McKendry, A., Brown-Guedira, G., et al. (2018). The accuracy of genomic prediction between environments and populations for soft wheat traits. *Crop Sci.* 58:2274. doi: 10.2135/cropsci2017.10.0638
- Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127, 595–607. doi: 10.1007/s00122-013-2243-1
- Jiang, M., Zhang, C., Hussain, K., Li, N., Sun, Q., Qing, M., et al. (2012). Pyramiding resistance genes to northern leaf blight and head smut in maize. *Int. J. Agric. Biol.* 14, 430–434.
- Kamiak (2021). *High Performance Computing*. Washington State University. Available online at: <https://hpc.wsu.edu/> (accessed January 21, 2021).
- Kharouf, S. H., Hamzeh, S. H., and Al-Azmeh, M. F. (2021). Races identification of wheat rusts in Syria during the 2019 growing season. *Arab J. Plant Prot.* 39, 1–13. doi: 10.22268/AJPP-39.1.001013
- 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
- Line, R. F., and Qayoum, A. (1992). Virulence, aggressiveness, evolution and distribution of races of *Puccinia striiformis* (the cause of stripe rust of wheat) in North America, 1968–87. *Tech. Bull. U. S. A.* Available online at: <http://agris.fao.org/agris-search/search.do?recordID=US9304750> (accessed January 16, 2020).
- Liu, L., Yuan, C. Y., Wang, M. N., See, D. R., Zemetra, R. S., and Chen, X. M. (2019). QTL analysis of durable stripe rust resistance in the North American winter wheat cultivar Skiles. *Theor. Appl. Genet.* 132, 1677–1691. doi: 10.1007/s00122-019-03307-2
- Liu, W., Naruoka, Y., Miller, K., Garland-Campbell, K. A., and Carter, A. H. (2018). Characterizing and validating stripe rust resistance loci in US Pacific Northwest winter wheat accessions (*Triticum aestivum* L.) by genome-wide association and linkage mapping. *Plant Genome* 11:87. doi: 10.3835/plantgenome2017.10.0087
- Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016a). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12:1005767. doi: 10.1371/journal.pgen.1005767
- Liu, Y., Chen, L., Liu, Y., Dai, H., He, J., Kang, H., et al. (2016b). Marker assisted pyramiding of two brown planthopper resistance genes, Bph3 and Bph2 (t), into elite rice cultivars. *Rice* 9, 1–7. doi: 10.1186/s12284-016-0096-3
- Liu, Y., Qie, Y. M., Li, X., Wang, M. N., and Chen, X. M. (2020). Genome-wide mapping of quantitative trait loci conferring all-stage and high-temperature adult-plant resistance to stripe rust in spring wheat landrace PI 181410. *Int. J. Mol. Sci.* 21:478. doi: 10.3390/ijms21020478
- Lozada, D. N., and Carter, A. H. (2019). Accuracy of single and multi-trait genomic prediction models for grain yield in US Pacific Northwest winter wheat. *Crop Breed. Genet. Genom.* 1:e190012. doi: 10.20900/cbpg2190012

- Lozada, D. N., and Carter, A. H. (2020). Insights into the genetic architecture of phenotypic stability traits in winter wheat. *Agronomy* 10:368. doi: 10.3390/agronomy10030368
- Lozada, D. N., Mason, R. E., Babar, M. A., Carver, B. F., Guedira, G.-B., Merrill, K., et al. (2017). Association mapping reveals loci associated with multiple traits that affect grain yield and adaptation in soft winter wheat. *Euphytica* 213:222. doi: 10.1007/s10681-017-2005-2
- McGowan, M., Wang, J., Dong, H., Liu, X., Jia, Y., Wang, X., et al. (2020). Ideas in genomic selection with the potential to transform plant molecular breeding: a review. *Preprints* 2020:202010.0460.v2. doi: 10.20944/preprints202010.0460.v2
- Merrick, L. F., and Carter, A. H. (2021). Comparison of genomic selection models for exploring predictive ability of complex traits in breeding programs. *bioRxiv*. 15:440015. doi: 10.1101/2021.04.15.440015
- Meuwissen, T. H. E., 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
- Michel, S., Ametz, C., Gungor, H., Akgöl, B., Epure, D., Grausgruber, H., et al. (2017). Genomic assisted selection for enhancing line breeding: merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theor. Appl. Genet.* 130, 363–376. doi: 10.1007/s00122-016-2818-8
- Michel, S., Ametz, C., Gungor, H., Epure, D., Grausgruber, H., Löschenberger, F., et al. (2016). Genomic selection across multiple breeding cycles in applied bread wheat breeding. *Theor. Appl. Genet.* 129, 1179–1189. doi: 10.1007/s00122-016-2694-2
- Milus, E. A., Lee, K. D., and Brown-Guedira, G. (2015). Characterization of stripe rust resistance in wheat lines with resistance gene *Yr17* and implications for evaluating resistance and virulence. *Phytopathology* 105, 1123–1130. doi: 10.1094/PHYTO-11-14-0304-R
- Mirdita, V., He, S., Zhao, Y., Korzun, V., Bothe, R., Ebmeyer, E., et al. (2015). Potential and limits of whole genome prediction of resistance to Fusarium head blight and Septoria tritici blotch in a vast Central European elite winter wheat population. *Theor. Appl. Genet.* 128, 2471–2481. doi: 10.1007/s00122-015-2602-1
- Mu, J., Liu, L., Liu, Y., Wang, M., See, D. R., Han, D., et al. (2020). Genome-wide association study and gene specific markers identified 51 genes or QTL for resistance to stripe rust in US winter wheat cultivars and breeding lines. *Front. Plant Sci.* 11:998. doi: 10.3389/fpls.2020.00998
- Muleta, K. T., Bulli, P., Zhang, Z., Chen, X., and Pumphrey, M. (2017). Unlocking diversity in germplasm collections via genomic selection: a case study based on quantitative adult plant resistance to stripe rust in spring wheat. *Plant Genome* 10:124. doi: 10.3835/plantgenome2016.12.0124
- Naruoka, Y., Garland-Campbell, K. A., and Carter, A. H. (2015). Genome-wide association mapping for stripe rust (*Puccinia striiformis* f. sp. tritici) in US Pacific Northwest winter wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 128, 1083–1101. doi: 10.1007/s00122-015-2492-2
- Ornella, L., Singh, S., Perez, P., Burgueño, J., Singh, R., Tapia, E., et al. (2012). Genomic prediction of genetic values for resistance to wheat rusts. *Plant Genome* 5, 136–148. doi: 10.3835/plantgenome2012.07.0017
- Peterson, R. F., Campbell, A. B., and Hannah, A. E. (1948). A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Can. J. Res.* 26c, 496–500. doi: 10.1139/cjr48c-033
- Pietrusińska, A., Czembor, J. H., and Czembor, P. C. (2011). Pyramiding two genes for leaf rust and powdery mildew resistance in common wheat. *Cereal Res. Commun.* 39, 577–588. doi: 10.1556/CRC.39.2011.4.13
- Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012). Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5, 103–113. doi: 10.3835/plantgenome2012.06.0006
- Poland, J., and Rutkoski, J. (2016). Advances and challenges in genomic selection for disease resistance. *Annu. Rev. Phytopathol.* 54, 79–98. doi: 10.1146/annurev-phyto-080615-100056
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <https://www.R-project.org/> (accessed April 28, 2021).
- Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., et al. (2016). Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor. Appl. Genet.* 129, 1843–1860. doi: 10.1007/s00122-016-2743-x
- Raymond, B., Bouwman, A. C., Schrooten, C., Houwing-Duistermaat, J., and Veerkamp, R. F. (2018). Utility of whole-genome sequence data for across-breed genomic prediction. *Genet. Sel. Evol.* 50, 1–12. doi: 10.1186/s12711-018-0396-8
- Rice, B., and Lipka, A. E. (2019). Evaluation of RR-BLUP genomic selection models that incorporate peak genome-wide association study signals in maize and sorghum. *Plant Genome* 12:180052. doi: 10.3835/plantgenome2018.07.0052
- Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., et al. (2015). Efficient use of historical data for genomic selection: a case study of stem rust resistance in wheat. *Plant Genome* 8:46. doi: 10.3835/plantgenome2014.09.0046
- Rutkoski, J. E., Heffner, E. L., and Sorrells, M. E. (2011). Genomic selection for durable stem rust resistance in wheat. *Euphytica* 179, 161–173. doi: 10.1007/s10681-010-0301-1
- Rutkoski, J. E., Poland, J. A., Singh, R. P., Huerta-Espino, J., Bhavani, S., Barbier, H., et al. (2014). Genomic selection for quantitative adult plant stem rust resistance in wheat. *Plant Genome* 7:6. doi: 10.3835/plantgenome2014.02.0006
- Schmidt, P., Hartung, J., Bennewitz, J., and Piepho, H.-P. (2019). Heritability in plant breeding on a genotype-difference basis. *Genetics* 212, 991–1008. doi: 10.1534/genetics.119.302134
- Singh, M., Mallick, N., Chand, S., Kumari, P., Sharma, J. B., Sivasamy, M., et al. (2017). Marker-assisted pyramiding of *Thinopyrum*-derived leaf rust resistance genes *Lr19* and *Lr24* in bread wheat variety HD2733. *J. Genet.* 96, 951–957. doi: 10.1007/s12041-017-0859-7
- Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redoña, E., Jannink, J.-L., et al. (2016). Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. *Heredity* 116, 395–408. doi: 10.1038/hdy.2015.113
- Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT Version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9:120. doi: 10.3835/plantgenome2015.11.0120
- Tekin, M., Cat, A., Akan, K., Catal, M., and Akar, T. (2021). A new virulent race of wheat stripe rust pathogen (*Puccinia striiformis* f. sp. tritici) on the resistance gene *Yr5* in Turkey. *Plant Dis.* doi: 10.1094/PDIS-03-21-0629-PDN
- Wang, D., Lin, Z., Kai, L. I., Ying, M. A., Wang, L., Yang, Y., et al. (2017). Marker-assisted pyramiding of soybean resistance genes *RSC4*, *RSC8*, and *RSC14Q* to soybean mosaic virus. *J. Integr. Agric.* 16, 2413–2420. doi: 10.1016/S2095-3119(17)61682-4
- Wang, M., and Chen, X. (2017). “Stripe rust resistance,” in *Stripe rust*, eds X. Chen and Z. Kang (Dordrecht: Springer), 353–558. doi: 10.1007/978-94-024-1111-9_5
- Wang, X.-Y., Chen, P. D., and Zhang, S.-Z. (2001). Pyramiding and marker-assisted selection for powdery mildew resistance genes in common wheat. *Yi Chuan Xue Bao* 28, 640–646.
- Ward, B. P., Brown-Guedira, G., Tyagi, P., Kolb, F. L., Van Sanford, D. A., Sneller, C. H., et al. (2019). Multi-environment and multitrait genomic selection models in unbalanced early-generation wheat yield trials. *Crop Sci.* 59:491. doi: 10.2135/cropsci2018.03.0189
- Wellings, C. R., Singh, R. P., Yahyaoui, A., Nazari, K., and McIntosh, R. A. (2009). “The development and application of near-isogenic lines for monitoring cereal rust pathogens,” in *Proceedings of Oral Papers and Posters, 2009 Technical Workshop, BGRI, Cd. Obregón, Sonora, Mexico*, 77–87.

- Wickham, H. (2011). ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* 3, 180–185. doi: 10.1002/wics.147
- Zhang, G. S., Zhao, Y. Y., Kang, Z. S., and Zhao, J. (2020). First report of a *Puccinia striiformis* f. sp. *tritici* race virulent to wheat stripe rust resistance gene *Yr5* in China. *Plant Dis.* 104:284. doi: 10.1094/PDIS-05-19-0901-PDN
- Zhang, Z., Ober, U., Erbe, M., Zhang, H., Gao, N., He, J., et al. (2014). Improving the accuracy of whole genome prediction for complex traits using the results of genome wide association studies. *PLoS ONE* 9:e93017. doi: 10.1371/journal.pone.0093017

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 Merrick, Burke, Chen and Carter. 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.



Genomic-Assisted Marker Development Suitable for CsCvy-1 Selection in Cucumber Breeding

OPEN ACCESS

Erdem Kahveci^{1†}, Zübeyir Devran^{2†}, Ercan Özkaynak^{3†}, Yiguo Hong^{4,5†}, David J. Studholme^{6†} and Mahmut Tör^{5*†}

Edited by:

Anna Maria Mastrangelo,
Council for Agricultural and
Economics Research (CREA), Italy

Reviewed by:

Shivendra Kumar,
Iowa State University, United States
Špela Baebler,
National Institute of Biology
(NIB), Slovenia

*Correspondence:

Mahmut Tör
m.tor@worc.ac.uk

†ORCID:

Erdem Kahveci
orcid.org/0000-0002-4467-5400
Zübeyir Devran
orcid.org/0000-0001-7150-284X
Ercan Özkaynak
orcid.org/0000-0002-4793-7963
Yiguo Hong
orcid.org/0000-0002-3352-9686
David J. Studholme
orcid.org/0000-0002-3010-6637
Mahmut Tör
orcid.org/0000-0002-4416-5048

Specialty section:

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

Received: 06 April 2021

Accepted: 22 June 2021

Published: 18 August 2021

Citation:

Kahveci E, Devran Z, Özkaynak E,
Hong Y, Studholme DJ and Tör M
(2021) Genomic-Assisted Marker
Development Suitable for CsCvy-1
Selection in Cucumber Breeding.
Front. Plant Sci. 12:691576.
doi: 10.3389/fpls.2021.691576

¹ M.Y. Genetik Tarım Tek. Lab. Tic. Ltd. Sti., Antalya, Turkey, ² Department of Plant Protection, Faculty of Agriculture, University of Akdeniz, Antalya, Turkey, ³ Yüksel Tohum Tarım San. ve Tic. A. S., Antalya, Turkey, ⁴ Research Centre for Plant RNA Signaling, College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, China, ⁵ Department of Biology, School of Science and the Environment, University of Worcester, Worcester, United Kingdom, ⁶ Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

Cucumber is a widely grown vegetable crop plant and a host to many different plant pathogens. *Cucumber vein yellowing virus* (CVYV) causes economic losses on cucumber crops in Mediterranean countries and in some part of India such as West Bengal and in African countries such as Sudan. CVYV is an RNA potyvirus transmitted mechanically and by whitefly (*Bemisia tabaci*) in a semipersistent manner. Control of this virus is heavily dependent on the management of the insect vector and breeding virus-resistant lines. DNA markers have been used widely in conventional plant breeding programs via marker-assisted selection (MAS). However, very few resistance sources against CVYV in cucumber exist, and also the lack of tightly linked molecular markers to these sources restricts the rapid generation of resistant lines. In this work, we used genomics coupled with the bulked segregant analysis method and generated the MAS-friendly Kompetitive allele specific PCR (KASP) markers suitable for CsCvy-1 selection in cucumber breeding using a segregating F₂ mapping population and commercial plant lines. Variant analysis was performed to generate single-nucleotide polymorphism (SNP)-based markers for mapping the population and genotyping the commercial lines. We fine-mapped the region by generating new markers down to 101 kb with eight genes. We provided SNP data for this interval, which could be useful for breeding programs and cloning the candidate genes.

Keywords: CVYV, cucumber, marker assisted selection, kompetitive allele-specific PCR genotyping, plant breeding

INTRODUCTION

Cucumber plants, *Cucumis sativus*, have been cultivated as a vegetable crop across the globe for centuries (Tatlioglu, 1993). The fruit is consumed as fresh or industrialized product, and the major producing countries are China (7,033,8971 tons), Turkey (1,916,645 tons), Russia (1,626,360 tons), Ukraine (1,034,170 tons), and Iran [871,692 tons (FAO, 2019)]. As an important vegetable, cucumber is challenged by many different fungal, oomycete, bacterial, and viral pathogens (Kong et al., 2015; Słomnicka et al., 2018; Bandamaravuri et al., 2020).

One of the most devastating viral pathogens is *Cucumber vein yellowing virus* (CVYV), which belongs to the *Potyviridae* family (Lecoq et al., 2000), has an RNA genome (Janssen et al., 2005), is transmitted mechanically and by whitefly, *Bemisia tabaci*, in a semipersistent manner (Mansour and Al-Musa, 1993), and infects a number of cucurbit species (Gil-Salas et al., 2011). The occurrence and heavy crop losses due to CVYV infection in the open fields and under protected cucumber crops have been reported in the Mediterranean countries from Israel to Portugal (Cohen and Nitzany, 1960; Louro et al., 2004). The main symptoms of CVYV on the cucumber include vein clearing followed by vein yellowing on the youngest leaves (Cohen and Nitzany, 1960), the occasional occurrence of yellow/green mosaics on the fruit (Cuadrado et al., 2007), and eventual general necrosis of the entire infected plant (Cohen and Nitzany, 1960). Mechanical transmission of the virus allows the use of cucumber as a test and indicative plant for multiplication.

Cucumber vein yellowing virus has been classified as a quarantine viral pathogen in the EPPO A2 Action List (https://www.eppo.int/ACTIVITIES/plant_quarantine/A2_list). Control of this virus relies heavily on the application of integrated pest management (IPM) practices that incorporate the ecosystem-based strategies, including cultural practices, biological and chemical control of the vector, and the use of resistant varieties (Horowitz et al., 2011). Sanitation, use of certified virus-free seedlings, and eradicating diseased plants parts are common practices for controlling viral plant pathogens (Hilje et al., 2001; Nazarov et al., 2020). Although chemical pesticides have been used to control the whitefly insect vector, concerns to human health, occurrence of insecticide resistance, and damage to the environment led to a search for alternative measures (Sani et al., 2020). Use of microbial biological control agents (MBCA), such as entomopathogenic fungi (Faria and Wraight, 2001; Sani et al., 2020), use of barrier or trap crops (Zhang et al., 2020), and use of beneficial insects, such as predators or parasitoids (Moreno-Ripoll et al., 2014), have been considered.

The RNA-guided genome editing using clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 has been also used to generate virus-resistant crops (Liu and Fan, 2014). For example, Chandrasekaran et al. (2016) used Cas9/sgRNA constructs to target the recessive *eukaryotic translation initiation factor 4E* (*eIF4E*) gene in cucumber. They reported that the homozygous T3 lines showed immunity to CVYV (Chandrasekaran et al., 2016) indicating the possibility of alternative new methods for CVYV control. Planting cultivars resistant to the whitefly and/or to the virus is one of the most important control measures in the CVYV management. In a study to identify cucumber lines resistant to whitefly, Novaes et al. (2020) screened 60 genotypes and found that accessions IAC-1214, IAC-1201, Campeiro, Japonês, IAC-1311, Kyria, and IAC-1175 displayed some low levels of attractiveness to these insects and suggested they could be included in the breeding programs to develop whitefly-resistant cucumber lines.

Genetics of resistance to CVYV have been investigated by several groups (Picó et al., 2003). A Spanish landrace of short

cucumber, C. sat-10, was found to be monogenic and displaying dominant resistance to CVYV (Picó et al., 2008). Similarly, a cucumber cultivar named Kyoto-3-feet originating from Japan has been reported to be resistant to CVYV (Martín-Hernández and Picó, 2021); however, detailed information on the nature of these resistance mechanisms is not available. Cucumber hybrid lines resistant to CVYV exist in the commercial market; however, currently all the work for selecting resistant lines relies on traditional pathotesting efforts. Recently, Pujol et al. (2019) described their elegant study on the resistant accession CE0749, a CVYV-resistant long Dutch-type cucumber. They used genomics and bulked segregant analysis (BSA) (Michelmore et al., 1991) and fine-mapped a locus containing the gene *CsCvy-1* locus in a 625 kb region with 24 candidate genes (Pujol et al., 2019).

Here, we described our investigations on the identification of DNA markers for fine mapping *CsCvy-1* using genomics and BSA. We used both segregating F₂ populations and the available commercial F₁ hybrids, mapped the locus down to 101 kb with eight genes, and provided single-nucleotide polymorphism (SNP) data for the interval, which could be useful for plant breeding programs.

MATERIALS AND METHODS

Plant Lines and Mapping Populations

An F₂ mapping population, generated from a cross between a susceptible (YT-189-1) and a resistant (YT-MLN-33) cucumber inbred lines (Yüksel Tohum A.S., Antalya, Turkey), was used in the phenotyping and genotyping experiments. F₃ families were raised by selfing the selected lines and used to determine the genotype of the F₂ lines.

Virus Isolate and Pathology Methods

The CVYV isolate used in this study was obtained from DSMZ (Braunschweig, Germany) and propagated in susceptible cucumber plants (*Cucumis sativus*, line YT-189-1). Virus inoculum was prepared by homogenizing 1 g infected leaves in 4 ml 0.01 M phosphate buffer (pH 7.0) containing 0.2% sodium sulfate and 0.2% diethyldithiocarbamic acid (DIECA, Sigma-Aldrich, St. Louis, MO, United States). After adding 600-mesh carborundum and active carbon, cotyledons of cucumber plants (parental lines, F₁, and F₃ generations), which were at the cotyledon to one-true-leaf stages, were mechanically inoculated. A second inoculation was performed 3 days after the first one to eliminate escapees. The inoculated cucumber seedlings were then kept in a growth chamber with temperature control set at 30/25°C (day/night) with a 16 h light/8 h dark photoperiod for 3 weeks and observed every 2 days. First symptoms were observed 5–7 days postinoculation (dpi), but became obvious after 12–15 dpi. After 3 weeks, no further symptom developments were observed; thus, 15 dpi was selected to be the optimal time for symptom evaluation. Plants showing clear symptoms including mosaic and vein yellowing on leaves were rated as susceptible, whereas those with no symptoms or a very light vein discoloration on only the oldest ones were accepted as resistant. A minimum of 20 plants was used per treatment.

DNA Extraction and Genome Sequencing

Young leaves were collected from parental and F₂ lines. Plant genomic DNA was isolated using the Wizard Magnetic Kit (Promega, Madison, WI, United States) following the instructions of manufacturer. DNA was extracted from each individual plant lines, and a gel electrophoresis was performed to determine whether high molecular weight DNAs were isolated. The resistant and susceptible bulks were generated from 20 resistant and 20 susceptible F₂ individuals, respectively, as described in earlier studies (Devran et al., 2015, 2018). Genomic DNA library and sequencing have been carried out by the University of Exeter Sequencing Service after quality check of DNAs, generating 2 × 150 bp paired-end read data for each parent line and bulked (resistant and susceptible) pools with Illumina HiSeq 2500 (Illumina, Inc. San Diego, CA, USA).

Analysis of Genomic Sequences

As previously described (Devran et al., 2018), we took the NGS analysis approach where the raw reads were trimmed using BBDuk (filter = 27, trimk = 27; <https://sourceforge.net/projects/bbmap/>) to remove Illumina adapters and to quality trim both ends to Q12. Subsequently, trimmed sequences from parental lines and the bulks were mapped onto the available reference cucumber genome (V2 and V3) using BBMap (<https://sourceforge.net/projects/bbmap/>), and the alignment data were converted to the BAM format (Li et al., 2009). As the CsCvy-1 locus was previously mapped onto chromosome 5 (Pujol et al., 2019), the data from the interval on chromosome 5: 7,000,000–7,850,000 were extracted using SAMtools (Li et al., 2009). The variant detection has been performed using BCFtools (Li et al., 2009) and a publicly available custom script (<https://github.com/davidstudholme/SNPsFromPileups>) as previously described; (Yemataw et al., 2018). Integrative Genomics Viewer (IGV) was used to visualize the alignment results (Robinson et al., 2011).

Converting Single Nucleotide Variants to PCR-Based Markers

Several of the SNPs within the interval were converted to Kompetitive Allele Specific PCR markers (KASP) by taking 100 bases either side of the SNP. KASP primers were developed using the LGC's primer picker software, Middlesex, United Kingdom. The PCRs were performed in a total volume of 15 µl that included DNA (10 ng; 5 µl), KASP Assay Mix (0.2 µl), KASP Master Mix (7.5 µl), and distilled water (2.3 µl). The KASP assay reactions were performed using the LightCycler[®] 480 II (Roche) using 61–55°C touchdown protocol (<https://biosearch-cdn.azureedge.net/assetsv6/KASP-thermal-cycling-conditions-all-protocols.pdf>). The fluorescence signal was measured for 2 min at 25°C using a FluOstar Omega Microplate Reader (BMG LABTECH, Ortenberg, Germany).

Confirming Interval and Identifying Marker-Assisted Selection (MAS)-Friendly Markers

As the CsCvy-1 locus had been previously mapped (Pujol et al., 2019), we used some of the published KASP markers including CVYV-184, CVYV-187, CVYV-188, CVYV-190, and

CVYV-122 in this work. Published and newly generated KASP markers were first tested on parents to confirm the identified polymorphisms and then 120 segregating F₂ lines. Marker genotyping data and the viral disease phenotyping data were used to confirm the CsCvy-1 interval. As we developed new markers (**Supplementary Table 1**) to narrow the interval down, we also tested these markers with the commercial F₁ hybrid lines, which were obtained from the relevant companies. As F₂ lines are a segregating population, markers discovered using F₂s may not be a reliable MAS-friendly marker. Therefore, we used F₁ hybrid lines to narrow the interval further down and identify the MAS-friendly markers.

Genomic Sequences and Accession Numbers

Cucumber reference genome sequences ChineseLong 9930 v2 are at <http://cucurbitgenomics.org/organism/2> ChineseLong 9930 v3 at (https://ftp.ncbi.nlm.nih.gov/genomes/genbank/plant/Cucumis_sativus/latest_assembly_versions/GCA_000004075.3_Cucumber_9930_V3/). The raw sequence reads aligning to the interval have been deposited in the Sequence Read Archive (SRA) and are accessible via BioProject accession PRJNA713378.

RESULTS

Resistance to CVYV Segregates as a Single Locus

A cross was generated between the susceptible *C. sativus* inbred line YT-189-1 and the resistant inbred line, YT-MLN-33. The F₁ hybrid showed resistance to CVYV, indicating that resistance was dominant. The F₁ was selfed to generate segregating F₂ populations. A total of 120 F₂ lines were taken to F₃ level, and 20 F₃ lines descending from each F₂s were inoculated with the virus to determine accurately the phenotype of the mapping population. Disease symptoms, including mosaics and vein yellowing, were obvious on the leaves of susceptible plants at 15 dpi (**Figure 1**). The segregation ratio observed in this bioassay was 92:28 (resistant:susceptible, 3:1; with Chi-square = 0.05 and $p \leq 0.05$), suggesting that a single locus was providing resistance to CVYV in this cross and allowing the subsequent analysis.

Linkage to CsCvy-1 Locus

We used a next-generation sequencing (NGS)-based BSA approach whereby we generated bulks from DNA isolated from 20 resistant and susceptible F₂ lines. We generated 150-bp paired-end Illumina HiSeq2500 sequencing data from the two parents and bulks (resistant and susceptible). A total of 390 million reads for each parent and 391 million reads for each bulk were generated. We then mapped these reads onto to the cucumber reference genome sequence (GenBank: GCA_000004075.3). However, during the course of our work, a locus designated CsCvy-1 mapped on chromosome 5 was published using BSA approach (Pujol et al., 2019). This prompted us to check whether we were mapping the same region even though we were using different breeding lines. We used published CVYV-184, CVYV-187, CVYV-188, CVYV-190, and CVYV-122 KASP markers (Pujol et al., 2019) to

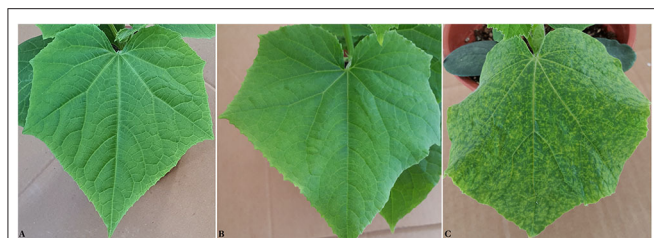


FIGURE 1 | Healthy and *Cucumber vein yellowing virus* (CVYV)-infected cucumber leaves. Cucumber plants were mechanically inoculated at the cotyledon to one-true-leaf stages. A second inoculation was performed 3 days after the first one. The inoculated cucumber seedlings were then kept in a growth chamber with temperature control set at 30/25°C (day/night) with a 16 h light/8 h dark photoperiod for 3 weeks and observed every 2 days. The plants were evaluated for symptom development at 15 days after the first inoculation (dpi). Control plants were treated with buffer without virus in a similar manner. **(A)** Leaf of an uninoculated control plant, **(B)** leaf of a *Cucumber vein yellowing virus* (CVYV)-inoculated-resistant plant, and **(C)** leaf of a CVYV-inoculated-susceptible plant.

TABLE 1 | Molecular markers used to define the interval for CsCvy-1 locus and the critical recombinant F₂ lines.

| F ₂ lines* | 10218 | 10317 | CsCvy-1 | 10644 | 10950 | 122** |
|-----------------------|-----------|-----------|---------|-----------|-----------|-----------|
| 58 | Rr | Rr | Rr | SS | SS | Rr |
| 59 | RR | RR | RR | Rr | Rr | Rr |
| 65 | Rr | Rr | Rr | RR | RR | Rr |
| 80 | Rr | RR | RR | RR | RR | RR |
| 103 | Rr | Rr | RR | RR | RR | RR |

*F₂ lines were generated from the cross between the resistant and the susceptible cultivars. **SS**, homozygous for susceptible parent allele; **RR**, homozygous for resistant parent allele; **RS**, heterozygous. Recombinants are shown in bold.

**This marker is from Pujol et al. (2019).

determine whether the resistance locus in our parental line YT-MLN-33 is linked to *CsCvy-1*. Our mapping data showed a clear linkage (Table 1, Supplementary Figure 1), and therefore we concentrated on chromosome 5. As we had already performed an SNP analysis using the then-available version of the reference genome sequence (GCA_000004075.2), we developed several KASP markers and mapped the *CsCvy-1* locus in our segregating mapping population. To make our work comparable with the published data, we then mapped our clean NGS reads onto the updated reference genome sequence (GCA_000004075.3), concentrated on a region between SNP10218, identified in this work, and the published CVYV122 marker (Supplementary Table 2). Several of the published markers were not polymorphic for the parental lines we used, e.g., CVYV-173, CVYV-174, CVYV-175, and CVYV 176.

Narrowing the Interval Using Nucleotide Variants

As our ultimate aim was to identify a marker that is tightly linked to *CsCvy-1*, we wanted to narrow the interval and generate further markers to identify an MAS-friendly marker. Using NGS data from parents and bulks, we mined the data

on chromosome 5: 10,218,000–11,370,000 (ChineseLong 9930 ASM407v2, Supplementary Figure 1). KASP markers were then designed and used for mapping to narrow the interval. A total of 13 new KASP markers were generated, and the locus was mapped down to a 327-kb interval between the markers 10,317 and 10,644 K using the available F₂ lines (Supplementary Table 2). As the version-three reference genome sequence became available, we also used this version and mined the data on chromosome 5: 7,000,000–7,850,000 (GCA_000004075.3, Supplementary Figure 2) for SNPs. A total of 436 SNPs have been detected (Supplementary Table 3). It should be noted here that although markers developed in this work using the GCA_000004075.V2 reference map to the region, several of the newly developed ones, especially toward the marker CVYV-122, were missing when the GCA_000004075.V3 reference was used. This may have been due to misassembly of the region as there was a 394 kb was missing in the GCA_000004075.3 genome (Figure 2).

Commercial Varieties Help Narrowing the Interval

Although we had enough number of markers to map the locus further, the number of F₂ lines to bring the interval down was not sufficient to identify further recombinants. We then obtained seeds of more than 20 commercial cucumber varieties with claimed CVYV phenotype and confirmed their phenotype by testing them with the CVYV isolate. Their DNAs were screened with our newly developed markers, and we narrowed down the locus to a 101-kb interval between the markers 10,317 and 10,418 K (Supplementary Table 4, Table 2). This finding suggests that the polymorphism identified in this work has been maintained across different varieties that have been used in the commercial breeding programs. In addition, the identified polymorphisms within the interval could be used in a breeding program by checking the existence of polymorphisms in the lines used.

The CsCvy-1 Interval Contains Genes That May Play a Role in Defense

CsCvy-1 locus mapped by Pujol et al. (2019) contained 24 genes. However, as we mapped the interval down to 101 kb, we used the annotations of the cucumber reference genome (GCA_000004075.3) to identify genes within the interval. The *CsCvy-1* locus in our mapping interval contains eight predicted genes (Table 3). Although *CsaV3_5G011160* encodes a cytochrome P450-like protein, *CsaV3_5G011170*, *CsaV3_5G011190*, and *CsaV3_5G011230* encode unknown proteins. However, *CsaV3_5G011180* encodes a serine/arginine repetitive matrix protein 2 isoform X2, *CsaV3_5G011220* encodes an endo-1,4-beta-xylanase, and two genes, *CsaV3_5G011200* and *CsaV3_5G011210*, both encode RNA-dependent RNA polymerase 1-like (RDR1-like) proteins. Interestingly, the deletion in the intragenic region of the RDR1 reported in this interval (Pujol et al., 2019) has been maintained in the resistant inbred line we used.

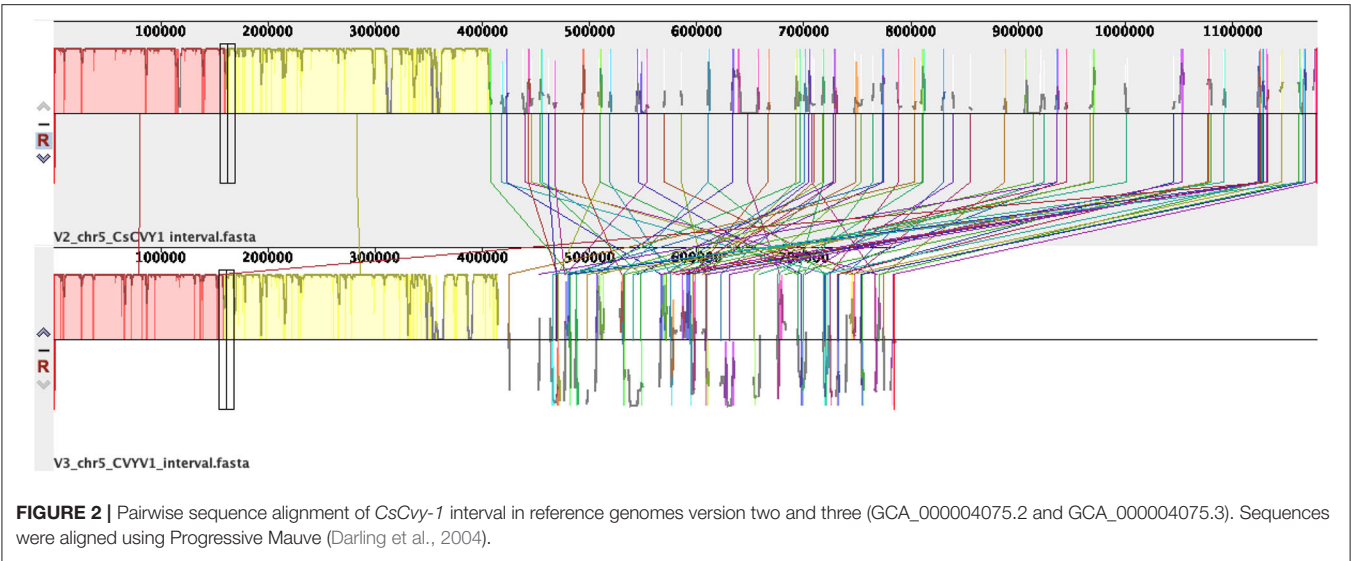


TABLE 2 | Molecular markers used to define the interval for *CsCvy-1* locus using F₁ hybrids.

| F ₁ hybrids* | 10270 | 10290 | 10317 | <i>CsCvy-1</i> ** | 10418 | 184*** | 10644 | 10950 | 122*** |
|-------------------------|-----------|-----------|-----------|-------------------|-------|--------|-----------|-----------|-----------|
| Civan | Rr | Rr | Rr | R**** | SS | SS | SS | SS | SS |
| Quinton | SS | SS | SS | SS | SS | SS | RR | SS | SS |
| Botanik | Rr | Rr | Rr | SS | SS | SS | RR | Rr | Rr |
| Kitir | SS | SS | SS | SS | SS | SS | RR | SS | SS |
| 52-23 | SS | SS | SS | SS | SS | SS | RR | SS | SS |
| Quarto | SS | SS | SS | SS | SS | SS | RR | SS | SS |
| Kybele | Rr | Rr | Rr | SS | SS | SS | RR | SS | SS |

*These were selected from more than 21 readily available varieties on the market.
**Phenotype information has been obtained from the web sites of companies, which sell these varieties to growers and confirmed by pathotesting.
***These markers were from Pujol et al. (2019).
****The phenotype was not confirmed by using offspring to determine whether it is homozygous or heterozygous.
RR, Homozygous resistant; Rr, Heterozygous resistant; SS, Homozygous susceptible. Recombination is shown in bold.

DISCUSSION

Here, we present genetic evidence that a single dominant locus *CsCvy-1* confers resistance to CVYV infection in our inbred lines, consistent with a recent report (Pujol et al., 2019). Our genomic and molecular investigations using an F₂ mapping population and also commercial resistant and susceptible varieties enabled us to map this locus down to 101-kb interval in which eight genes reside. Use of genomics allowed the identification of SNPs that could be used in breeding programs.

The plant immune system has the ability to recognize extracellular or intracellular molecules derived from plant pathogens and generate a defense response to restrict the pathogen growth or replication (Wang et al., 2008; Tör et al., 2009; Steinbrenner et al., 2015). Map-based studies usually involve the phenotyping and genotyping of a large number of individual plants in a segregating population. Using this

TABLE 3 | Genes within the *CsCvy-1* interval in the Chinese Long cucumber genome.

| Gene ID | Putative function |
|-----------------------|--|
| <i>CsaV3_5G011160</i> | Cytochrome P450-like protein |
| <i>CsaV3_5G011170</i> | Unknown protein |
| <i>CsaV3_5G011180</i> | Serine/arginine repetitive matrix protein 2 isoform X2 |
| <i>CsaV3_5G011190</i> | Unknown protein |
| <i>CsaV3_5G011200</i> | RNA-dependent RNA polymerase 1-like |
| <i>CsaV3_5G011210</i> | RNA-dependent RNA polymerase 1-like |
| <i>CsaV3_5G011220</i> | Endo-1,4-beta-xylanase |
| <i>CsaV3_5G011230</i> | Unknown protein |

approach, genes conferring resistance to fungal, oomycete, bacterial, and viral pathogens have been mapped, and many of them have been cloned (Tai et al., 1999; Borhan et al., 2008; Kim et al., 2017; Chen et al., 2021).

Linkage analysis plays a significant role in the cloning genes or generating markers tightly linked to the locus of interest using map-based approach. When we started this investigation, there was no published data on the chromosomal location of the gene conferring resistance to CVYV. During our SNP analysis, Pujol et al. (2019) published their work on the mapping of *CsCvy-1* in cucumber using genomic approach. We used relevant markers from this published work; however, several of them were not polymorphic in our parental lines, indicating the importance of generating SNP data from the lines used in generating mapping populations. After establishing the linkage between the resistance source in our material and the *CsCvy-1*, it was obvious that we were mapping the same locus, and thus we zoomed into the region.

We used genomics and BSA previously to clone genes (Woods-Tör et al., 2018) and to generate MAS-friendly molecular markers (Devran et al., 2018) that are tightly linked to the gene of interest. Our experience shows that although the use of reference genomes helps the identification of variants in the region of interest, different versions of reference genome

assemblies produced different results in the SNP analysis. It was the case in this study where we initially used version 2 (GCA_000004075.2) as the reference and generated markers for our mapping work. Although all the markers generated from version 2 mapped the gene, several of them were missing when version 3 (GCA_000004075.3) were used, indicating the importance of mapping for confirmation and using more than one available reference genome.

High number of individual lines in a map-based study help identify the recombinant lines, which enables narrowing the interval. It can be easy to generate large number of F₂ lines from plants, such as *Arabidopsis thaliana* (Tör et al., 2002). However, in plants such as cucumber, it may not be possible to achieve large number of F₂s. In this work, we relied on 120 F₂ lines to narrow the interval down to a 327 kb. Considering the breeding efforts where many characters are collected in a “pure” line, during which many crosses are carried out and many recombination events take place, for an MAS-friendly marker, the interval needs to be very small so that the likelihood of a recombination event between the marker and the gene of interest is almost zero. Bearing this in mind, we used the commercial cucumber F₁ hybrids in our phenotyping and genotyping assays and reduced the interval down to 101 kb with eight genes.

Resistance to plant pathogens could be provided by membrane-bound proteins, such as receptor-like proteins (RLPs) (Wang et al., 2008) or receptor-like kinases (RLKs) (Roux et al., 2011; Zhang et al., 2013) or by the cytoplasmic nucleotide-binding, leucine-rich repeat (NLR) immune receptors (Adachi et al., 2019). There were no classic RLP, RLK, or NLR-type genes in the 101-kb CsCvy-1 interval. Pujol et al. (2019) looked into the small variants and structural variation in the locus and argued that *CsaV3_5G011180* encoding for serine/arginine repetitive matrix protein (SARMP) could be a possible candidate. In addition, Pujol et al. (2019) postulated that *CsaV3_5G011200* and *CsaV3_5G011210* encoding RDRs 1a and 1b had the most appealing modifications in the locus and discussed the role of RDRs in RNA silencing pathways. Leibman et al. (2018) carried out detailed investigations into the *RDR1*-like genes in cucumber and reported the presence of four putative *RDR1*-family genes. They then investigated the expression of these *RDR1*-like genes and their role in defense against different viruses, including *Zucchini yellow mosaic virus* (ZYMV), CMV, and CVYV and showed that the level of *RDR1*-like gene expression varied according to the virus used (Leibman et al., 2018).

The NLR-type disease resistance genes in *Arabidopsis* have been reported to be clustered in the genome (Holub, 2007), and some of them function together and could be in head-to-head orientation, termed paired NLRs (Saucet et al., 2015). Further detailed studies indicated that one of them could function as a pathogen sensor, and the other member as signaling executor (Van de Weyer et al., 2019). Here, we have *RDR1a* and *RDR1b* in the interval right next to each other, functioning “like an *R*-gene” (Leibman et al., 2018), but it is not totally clear from expression studies whether they function together as some genetic investigations are needed. It is tempting to speculate that *RDR1a* and *RDR1b* are the most suitable candidate genes for the CVYV resistance.

Our strategy to use genomics and BSA to identify SNPs and generate molecular markers that could be employed in the selection of CsCvy-1 enabled us to screen several markers and narrowed the interval down. These SNPs and markers could be used to identify polymorphism in different backgrounds in any breeding program to select CsCvy-1. Subsequent experiments could be designed to silence both *RDR1* and *RDR2* genes individually and together in the same background to reveal their dependence onto each other and their contribution to CVYV defense.

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 at: <https://www.ncbi.nlm.nih.gov/genbank/>, PRJNA713378.

AUTHOR CONTRIBUTIONS

EK, ZD, and MT planned and designed the research. EK carried out pathology tests. ZD performed genotyping using markers. EÖ made crossings between plant lines and produced offspring. YH, MT, and DS analyzed and interpreted the data. EK, ZD, YH, DS, and MT wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors declare that this study received funding from M.Y. Genetik Tarım Tek. Lab. Tic. Ltd. Sti. and Yüksel Tohum A.S. The funder had the following involvement in the study: M.Y. Genetik Tarım Tek. Lab. Tic. Ltd. Sti. provided equipments to carry out pathology and molecular tests, and Yüksel Tohum A.S. provided seeds and plant materials. The sequencing work in this project utilized equipment funded by the Wellcome Trust Institutional Strategic Support Fund (WT097835MF), Wellcome Trust Multi User Equipment Award (WT101650MA), and BBSRC LOLA award (BB/K003240/1).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.691576/full#supplementary-material>

Supplemental Figure 1 | Position of SNP-based markers on chromosome 5 in ChineseLong 9930 ASM407v2.

Supplemental Figure 2 | Position of SNP-based markers on chromosome 5 in ChineseLong 9930 GCA_000004075.3.

Supplemental Table 1 | KASP marker sequences developed in this study and previously published work.

Supplemental Table 2 | Mapping CsCvy-1 with existing and newly generated KASP markers using F₂ population.

Supplemental Table 3 | SNPs and their positions in CsCvy-1 locus on chromosome 5 using GCA_000004075.3. Proportions of aligned reads supporting the variant base are given.

Supplemental Table 4 | Mapping CsCvy-1 with existing and newly generated KASP markers using F₁ hybrids.

REFERENCES

- Adachi, H., Contreras, M. P., Harant, A., Wu, C. H., Derevnina, L., Sakai, T., et al. (2019). An N-terminal motif in NLR immune receptors is functionally conserved across distantly related plant species. *Elife* 8:121. doi: 10.7554/eLife.49956
- Bandamaravuri, K. B., Nayak, A. K., Bandamaravuri, A. S., and Samad, A. (2020). Simultaneous detection of downy mildew and powdery mildew pathogens on *Cucumis sativus* and other cucurbits using duplex-qPCR and HRM analysis. *AMB Express* 10:135. doi: 10.1186/s13568-020-01071-x
- Borhan, M. H., Gunn, N., Cooper, A., Gulden, S., Tör, M., Rimmer, S. R., et al. (2008). WRR4 encodes a TIR-NB-LRR protein that confers broad-spectrum white rust resistance in *Arabidopsis thaliana* to four physiological races of *Albugo candida*. *MPMI* 21, 757–768. doi: 10.1094/MPMI-21-6-0757
- Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., et al. (2016). Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol. Plant Pathol.* 17, 1140–1153. doi: 10.1111/mpp.12375
- Chen, C., Jost, M., Clark, B., Martin, M., Matny, O., Steffenson, B. J., et al. (2021). BED-domain-containing NLR from wild barley confers resistance to leaf rust. *Plant Biotechnol. J.* 19:1206–1215. doi: 10.1111/pbi.13542
- Cohen, S., and Nitzany, F. (1960). A whitefly transmitted virus of Cucurbits in Israel. *Phytopathol. Mediterranea* 1, 44–46.
- Cuadrado, I. M., Janssen, D., Velasco, L., Ruiz, L., and Segundo, E. (2007). First report of cucumber vein yellowing virus in Spain. *Plant Dis.* 85, 336–336. doi: 10.1094/PDIS.2001.85.3.336A
- Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704
- Devran, Z., Kahveci, E., Hong, Y., Studholme, D. J., and Tör, M. (2018). Identifying molecular markers suitable for *Frl* selection in tomato breeding. *Theor. Appl. Genet.* 131: 2099–2105. doi: 10.1007/s00122-018-3136-0
- Devran, Z., Kahveci, E., Özkaynak, E., Studholme, D. J., and Tör, M. (2015). Development of molecular markers tightly linked to *Pvr4* gene in pepper using next-generation sequencing. *Mol. Breed.* 35:101. doi: 10.1007/s11032-015-0294-5
- FAO (2019). Available online at: <http://www.fao.org/faostat/en/#data/QC> (accessed June 5, 2021).
- Faria, M., and Wraight, S. P. (2001). Biological control of *Bemisia tabaci* with fungi. *Crop Prot.* 20, 767–778. doi: 10.1016/S0261-2194(01)00110-7
- Gil-Salas, F. M., Peters, J., Boonham, N., Cuadrado, I. M., and Janssen, D. (2011). Yellowing disease in zucchini squash produced by mixed infections of Cucurbit yellow stunting disorder virus and Cucumber vein yellowing virus. *Phytopathology* 101, 1365–1372. doi: 10.1094/PHYTO-12-10-0343
- Hilje, L., Costa, H. S., and Stansly, P. A. (2001). Cultural practices for managing *Bemisia tabaci* and associated viral diseases. *Crop Prot.* 20, 801–812. doi: 10.1016/S0261-2194(01)00112-0
- Holub, E. B. (2007). Natural variation in innate immunity of a pioneer species. *Curr. Opin. Plant Biol.* 10, 415–424. doi: 10.1016/j.pbi.2007.05.003
- Horowitz, A. R., Antignus, Y., and Gerling, D. (2011). “Management of *Bemisia tabaci* whiteflies,” in *The Whitefly, Bemisia tabaci* (Homoptera: Aleyrodidae) Interaction With Geminivirus-Infected Host Plants, ed W. Thompson (Dordrecht: Springer), 293–322. doi: 10.1007/978-94-007-1524-0_11
- Janssen, D., Martin, G., Velasco, L., Gómez, P., Segundo, E., Ruiz, L., et al. (2005). Absence of a coding region for the helper component-proteinase in the genome of cucumber vein yellowing virus, a whitefly-transmitted member of the Potyviridae. *Arch. Virol.* 150, 1439–1447. doi: 10.1007/s00705-005-0515-z
- Kim, S. B., Kang, W. H., Huy, H. N., Yeom, S. I., An, J. T., Kim, S., et al. (2017). Divergent evolution of multiple virus-resistance genes from a progenitor in *Capsicum* spp. *New Phytol.* 213: 886–899. doi: 10.1111/nph.14177
- Kong, W., Chen, N., Liu, T., Zhu, J., Wang, J., He, X., et al. (2015). Large-scale transcriptome analysis of cucumber and *Botrytis cinerea* during infection. *PLoS ONE* 10:e0142221. doi: 10.1371/journal.pone.0142221
- Lecoq, H., Desbiez, C., Delécolle, B., Cohen, S., and Mansour, A. (2000). Cytological and molecular evidence that the whitefly-transmitted Cucumber vein yellowing virus is a tentative member of the family Potyviridae. *J. General Virol.* 81, 2289–2293. doi: 10.1099/0022-1317-81-9-2289
- Leibman, D., Kravchik, M., Wolf, D., Haviv, S., Weissberg, M., Ophir, R., et al. (2018). Differential expression of cucumber RNA-dependent RNA polymerase 1 genes during antiviral defence and resistance. *Mol. Plant Pathol.* 19, 300–312. doi: 10.1111/mpp.12518
- 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
- Liu, L., and Fan, X. D. (2014). CRISPR–Cas system: a powerful tool for genome engineering. *Plant Mol. Biol.* 85, 209–218. doi: 10.1007/s11103-014-0188-7
- Louro, D., Quinot, A., Neto, E., Fernandes, J. E., Marian, D., Vecchiati, M., et al. (2004). Occurrence of Cucumber vein yellowing virus in cucurbitaceous species in southern Portugal. *Plant Pathol.* 53, 241–241. doi: 10.1111/j.0032-0862.2004.00996.x
- Mansour, A., and Al-Musa, A. (1993). Cucumber vein yellowing virus; host range and virus vector relationships. *J. Phytopathol.* 137, 73–78. doi: 10.1111/j.1439-0434.1993.tb01327.x
- Martín-Hernández, A. M., and Picó, B. (2021). Natural resistances to viruses in cucurbits. *Agronomy* 11:23. doi: 10.3390/agronomy11010023
- Michelmores, R. W., Paran, I., and Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U.S.A.* 88, 9828–9832. doi: 10.1073/pnas.88.21.9828
- Moreno-Ripoll, R., Gabarra, R., Symondson, W. O. C., King, R. A., and Agustí N. (2014). Do the interactions among natural enemies compromise the biological control of the whitefly *Bemisia tabaci*? *J. Pest Sci.* 87, 133–141. doi: 10.1007/s10340-013-0522-x
- Nazarov, P. A., Baleev, D. N., Ivanova, M. I., Sokolova, L. M., and Karakozova, M. V. (2020). Infectious plant diseases: etiology, current status, problems and prospects in plant protection. *Acta Naturae* 12, 46–59. doi: 10.32607/actanaturae.11026
- Novaes, N. S., Lourenço, A. L., Bentivenha, J. P. F., Baldin, E. L. L., and Melo, A. M. T. (2020). Characterization and potential mechanisms of resistance of cucumber genotypes to *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Phytoparasitica* 48, 643–657. doi: 10.1007/s12600-020-00826-3
- Picó, B., Sifres, A., Martínez-Pérez, E., Leiva-Brondo, M., and Nuez, F. (2008). “Genetics of the resistance to CVYV in cucumber. In Modern variety breeding for present and future needs,” in *Proceedings of the 18th EUCARPIA General Congress*, eds J. Prohens and M. L. Badenes (Valencia), 452–456.
- Picó, B., Villar, C., Nuez, F., and Weber, W. E. (2003). Screening *Cucumis sativus* landraces for resistance to cucumber vein yellowing virus. *Plant Breeding* 122, 426–430. doi: 10.1046/j.1439-0523.2003.00882.x
- Pujol, M., Alexiou, K. G., Fontaine, A. S., Mayor, P., Miras, M., Jahrmann, T., et al. (2019). Mapping Cucumber vein yellowing virus resistance in cucumber (*Cucumis sativus* L.) by using BSA-seq analysis. *Front. Plant Sci.* 10:1583. doi: 10.3389/fpls.2019.01583
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754
- Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., et al. (2011). The *Arabidopsis* leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell* 23, 2440–2455. doi: 10.1105/tpc.111.084301
- Sani, I., Ismail, S. I., Abdullah, S., Jalinas, J., Jamian, S., and Saad, N. (2020). A review of the biology and control of whitefly, *Bemisia tabaci* (hemiptera: Aleyrodidae), with special reference to biological control using entomopathogenic fungi. *Insects* 11:619. doi: 10.3390/insects11090619
- Saucet, S. B., Ma, Y., Sarris, P. F., Furzer, O. J., Sohn, K. H., and Jones, J. D. (2015). Two linked pairs of *Arabidopsis* TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nat. Commun.* 6, 6338–6312. doi: 10.1038/ncomms7338
- Slomnicka, R., Olczak-Woltman, H., Korzeniewska, A., Gozdowski, D., Niemirówicz-Szczytt, K., and Bartoszewski, G. (2018). Genetic mapping of *psl* locus and quantitative trait loci for angular leaf spot resistance in cucumber (*Cucumis sativus* L.). *Mol. Breeding* 38, 111–119. doi: 10.1007/s11032-018-0866-2
- Steinbrenner, A. D., Goritschnig, S., and Staskawicz, B. J. (2015). Recognition and activation domains contribute to allele-specific responses of an *Arabidopsis*

- NLR receptor to an oomycete effector protein. *PLoS Pathog.* 11:e1004665. doi: 10.1371/journal.ppat.1004665
- Tai, T. H., Dahlbeck, D., and Clark, E. T., Gajiwala, P., Pasion, R., Whalen, M. C., et al. (1999). Expression of the *Bs2* pepper gene confers resistance to bacterial spot disease in tomato. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14153–14158. doi: 10.1073/pnas.96.24.14153
- Tatlioglu, T. (1993). “Cucumber: *Cucumis sativus* L.” in *Genetic Improvement of Vegetable Crops*, eds G. Kalloo and B. O. Bergh (Pergamon), 197–234. doi: 10.1016/B978-0-08-040826-2.50017-5
- Tör, M., Gordon, P., Cuzick, A., Eulgem, T., Sinapidou, E., Mert-Türk, F., et al. (2002). Arabidopsis *SGT1b* is required for defense signaling conferred by several downy mildew resistance genes. *Plant Cell* 14, 993–1003. doi: 10.1105/tpc.001123
- Tör, M., Lotze, M. T., and Holton, N. (2009). Receptor-mediated signalling in plants: molecular patterns and programmes. *J. Exp. Bot.* 60, 3645–3654. doi: 10.1093/jxb/erp233
- Van de Weyer, A. L., Monteiro, F., Furzer, O. J., Nishimura, M. T., Cevik, V., Witek, K., et al. (2019). A species-wide inventory of NLR genes and alleles in *Arabidopsis thaliana*. *Cell* 178, 1260–1272.e14. doi: 10.1016/j.cell.2019.07.038
- Wang, G., Ellendorff, U., Kemp, B., Mansfield, J. W., Forsyth, A., Mitchell, K., et al. (2008). A genome-wide functional investigation into the roles of receptor-like proteins in *Arabidopsis*. *Plant Physiol.* 147, 503–517. doi: 10.1104/pp.108.119487
- Woods-Tör, A., Studholme, D. J., Cevik, V., Telli, O., Holub, E. B., and Tör, M. (2018). A suppressor/avirulence gene combination in *Hyaloperonospora arabidopsidis* determines race specificity in *Arabidopsis thaliana*. *Front. Plant Sci.* 9:1957. doi: 10.3389/fpls.2018.00265
- Yemataw, Z., Muzemil, S., Ambachew, D., Tripathi, L., Tesfaye, K., Chala, A., Farbos, A., et al. (2018). Genome sequence data from 17 accessions of *Ensete ventricosum*, a staple food crop for millions in Ethiopia. *Data Brief* 18, 285–293. doi: 10.1016/j.dib.2018.03.026
- Zhang, W., Fraiture, M., Kolb, D., Löffelhardt, B., Desaki, Y., Boutrot, F. F., et al. (2013). Arabidopsis RECEPTOR-LIKE PROTEIN30 and Receptor-Like Kinase SUPPRESSOR OF BIR1-1/EVERSHED mediate innate immunity to necrotrophic fungi. *Plant Cell* 25, 4227–4241. doi: 10.1105/tpc.113.117010
- Zhang, X. M., Lövei, G. L., Ferrante, M., Yang, N. W., and Wan, F. H. (2020). The potential of trap and barrier cropping to decrease densities of the whitefly *Bemisia tabaci* MED on cotton in China. *Pest Manag. Sci.* 76, 366–374. doi: 10.1002/ps.5524

Conflict of Interest: EK was employed by company M.Y. Genetik Tarim Tek. Lab. Tic. Ltd. Sti. EÖ as employed by company Yüksel Tohum Tarim San. ve Tic. A. S.

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.

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 Kahveci, Devran, Özkaynak, Hong, Studholme and Tör. 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 Association Study Reveals Novel Genetic Loci for Quantitative Resistance to Septoria Tritici Blotch in Wheat (*Triticum aestivum* L.)

Tilahun Mekonnen¹, Clay H. Sneller², Teklehaimanot Haileselassie¹, Cathrine Ziyomo², Bekele G. Abeyo³, Stephen B. Goodwin⁴, Dagnachew Lule⁵ and Kassahun Tesfaye^{1,6*}

OPEN ACCESS

Edited by:

Harsh Raman,
New South Wales Department of
Primary Industries, Australia

Reviewed by:

Pasquale De Vita,
Research Centre for Industrial
Crops, Italy
Surinder Banga,
Punjab Agricultural University, India
Deepmala Sehgal,
International Maize and Wheat
Improvement Center, Mexico

*Correspondence:

Kassahun Tesfaye
kassahun.tesfaye@aau.edu.et;
kassahuntesfaye@yahoo.com

Specialty section:

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

Received: 23 February 2021

Accepted: 19 August 2021

Published: 24 September 2021

Citation:

Mekonnen T, Sneller CH,
Haileselassie T, Ziyomo C, Abeyo BG,
Goodwin SB, Lule D and Tesfaye K
(2021) Genome-Wide Association
Study Reveals Novel Genetic Loci for
Quantitative Resistance to Septoria
Tritici Blotch in Wheat (*Triticum*
aestivum L.).
Front. Plant Sci. 12:671323.
doi: 10.3389/fpls.2021.671323

¹ Institute of Biotechnology, Addis Ababa University, Addis Ababa, Ethiopia, ² Biosciences Eastern and Central Africa (BecA), Nairobi, Kenya, ³ International Maize and Wheat Improvement Center- CIMMYT (Ethiopia), Addis Ababa, Ethiopia, ⁴ United States Department of Agriculture (USDA)-Agricultural Research Service, West Lafayette, IN, United States, ⁵ Oromia Agricultural Research Institute (OARI), Addis Ababa, Ethiopia, ⁶ Ethiopian Biotechnology Institute (EBTI), Addis Ababa, Ethiopia

Septoria tritici blotch, caused by the fungus *Zymoseptoria tritici*, poses serious and persistent challenges to wheat cultivation in Ethiopia and worldwide. Deploying resistant cultivars is a major component of controlling septoria tritici blotch (STB). Thus, the objective of this study was to elucidate the genomic architecture of STB resistance in an association panel of 178 bread wheat genotypes. The association panel was phenotyped for STB resistance, phenology, yield, and yield-related traits in three locations for 2 years. The panel was also genotyped for single nucleotide polymorphism (SNP) markers using the genotyping-by-sequencing (GBS) method, and a total of 7,776 polymorphic SNPs were used in the subsequent analyses. Marker-trait associations were also computed using a genome association and prediction integrated tool (GAPIT). The study then found that the broad-sense heritability for STB resistance ranged from 0.58 to 0.97 and 0.72 to 0.81 at the individual and across-environment levels, respectively, indicating the presence of STB resistance alleles in the association panel. Population structure and principal component analyses detected two sub-groups with greater degrees of admixture. A linkage disequilibrium (LD) analysis in 338,125 marker pairs also detected the existence of significant ($p \leq 0.01$) linkage in 27.6% of the marker pairs. Specifically, in all chromosomes, the LD between SNPs declined within 2.26–105.62 Mbp, with an overall mean of 31.44 Mbp. Furthermore, the association analysis identified 53 loci that were significantly (false discovery rate, FDR, <0.05) associated with STB resistance, further pointing to 33 putative quantitative trait loci (QTLs). Most of these shared similar chromosomes with already published Septoria resistance genes, which were distributed across chromosomes 1B, 1D, 2A, 2B, 2D, 3A, 3B, 3D, 4A, 5A, 5B, 6A, 7A, 7B, and 7D. However, five of the putative QTLs identified on chromosomes 1A, 5D, and 6B appeared to be novel. Dissecting the detected loci on IWGSC RefSeq Annotation v2.1 revealed the existence of disease resistance-associated genes in the identified QTL regions that are involved in plant defense responses. These putative QTLs explained

2.7–13.2% of the total phenotypic variation. Seven of the QTLs ($R^2 = 2.7\text{--}10.8\%$) for STB resistance also co-localized with marker-trait associations (MTAs) for agronomic traits. Overall, this analysis reported on putative QTLs for adult plant resistance to STB and some important agronomic traits. The reported and novel QTLs have been identified previously, indicating the potential to improve STB resistance by pyramiding QTLs by marker-assisted selection.

Keywords: genome-wide association study, linkage disequilibrium, population structure, quantitative trait locus, septoria tritici blotch, wheat, *Zymoseptoria tritici*

INTRODUCTION

Common wheat (*Triticum aestivum* L.) is the most widely cultivated and the major staple food crop in the world consumed by human, providing almost 20% of the total calories and 21% of protein demand globally (Arzani and Ashraf, 2017; International Wheat Genome Sequencing Consortium (IWGSC), 2018; Ramadas et al., 2019). By 2050, the world's human population is projected to reach nine billion, and we will need to increase wheat production by 70% to feed this projected growth (FAO, 2009; Ray et al., 2013; Marcussen et al., 2014). Hence, boosting the wheat harvest is very pertinent to achieve zero-hunger by 2050.

Septoria tritici blotch, caused by the fungus *Zymoseptoria tritici* (anamorph: *Septoria tritici*), is an ever-existing bottleneck to wheat cultivation worldwide (Dalvand et al., 2018; Odilbekov et al., 2019), accounting for 30–54% of global wheat yield loss annually (Eyal and Levy, 1987). Septoria tritici blotch (STB) is also a major threat to wheat production in Ethiopia (Getinet et al., 1990; Takele et al., 2015; Kidane et al., 2017; Mekonnen et al., 2019, 2020), causing up to 82% of yield loss in the worst seasons (Getinet et al., 1990; Mengistu et al., 1991; Ayele et al., 2008).

The deployment of genetic resistance is the most durable, economical, and environmentally friendly method to manage crop diseases like STB (Ghaneie et al., 2011; Mekonnen et al., 2019; Odilbekov et al., 2019). In particular, qualitative and quantitative types of resistance to STB have been reported in wheat (Arraiano and Brown, 2006; Arraiano et al., 2009). The former refers to a condition where one or few major *Stb* genes provide resistance to specific *Z. tritici* isolates (Brown et al., 2015). Quantitative resistance, on the other hand, results from the expression of many genes with minor effects and is generally not specific to isolates. As such, quantitative resistance is the

most effective, durable, and preferred method to manage rapidly evolving wheat pathogens such as *Z. tritici* (Long et al., 2019).

The resistance-breeding method used in Ethiopia is mainly conventional, making the crop-improvement program very slow and inefficient. Nowadays, the advent and application of modern genomic tools have revolutionized crop breeding by facilitating the precise identification, mapping, and introgression of genomic regions controlling useful agronomic traits, such as resistance, into new cultivars. To account for this, a genome-wide association study (GWAS) is a powerful approach to elucidating the genomic architecture of many traits (Long et al., 2019). The development of high-throughput sequencing and bioinformatics technologies (Huang et al., 2017) has also enabled GWAS to scan single nucleotide polymorphisms (SNPs) associated with desirable traits at the whole-genome scale (Rafalski, 2010).

Genome-wide association studies have been successfully applied to many crop species (Xiao et al., 2017) such as maize (Rashid et al., 2018), rice (Huang et al., 2017), wheat (Kidane et al., 2017; Long et al., 2019; Cheng et al., 2020), and sorghum (Girma et al., 2019). In particular, this study design has been used in wheat to analyze several traits such as resistance to stripe rust (Long et al., 2019; Yao et al., 2019; Cheng et al., 2020), stem rust (Edae et al., 2015; Kankwatsa et al., 2017), Septoria tritici blotch (Kidane et al., 2017; Odilbekov et al., 2019), drought tolerance (Mathew et al., 2019), and other phenological characteristics, plus yield and yield-related traits (Jamil et al., 2019; Wang et al., 2019; Ward et al., 2019). While Ethiopia is the largest producer of wheat in sub-Saharan Africa, little is known about the resistance Ethiopian wheat cultivars have to STB, even though it is the most important disease economically. Thus, the objectives of this study were: (1) to determine the population structure, family relatedness, and level of linkage disequilibrium of the tested bread wheat association panel; (2) to elucidate the genomic architecture of adult plant resistance to STB; (3) to identify the SNP loci underlying yield, yield-related, and phenological traits in Ethiopian cultivars that could be useful in wheat breeding programs.

MATERIALS AND METHODS

Association Mapping Panel

This study used an association panel of 180 bread wheat (*Triticum aestivum* L.) genotypes (Supplementary Table 1), of which 167 were obtained from the International Maize and Wheat Improvement Center (CIMMYT-Mexico) and 13 were

Abbreviations: DH, days to heading; DF, days to flowering; DM, days to maturity; FDR, false discovery rate; FarmCPU, fixed and random model circulating probability unification; GAPIT, genome association and prediction integrated tools; GBS, genotyping by sequencing; GFD, grain-filling duration; GWAS, genome-wide association study; HLW, hectoliter weight; PH, plant height; LD, linkage disequilibrium; MAF, minor allele frequency; MAS, marker-assisted selection; MTAs, marker-trait associations; PCA, principal component analysis; QTL, quantitative trait locus or loci; SL, spike length; SN, seed number per spike; SDS, Septoria disease severity; SDSH, Septoria disease severity at heading; SDSMM, Septoria disease severity at mid-maturity; SDSM, Septoria disease severity at maturity; STB, Septoria tritici blotch; SNP, single-nucleotide polymorphism; SPC, Septoria progress coefficient; TKW, thousand kernel weight.

commercial cultivars grown in Ethiopia. The 167 CIMMYT genotypes included 49 from the International Bread Wheat Screening Nursery (IBWSN), 56 from the International Septoria Observation Nursery (ISEPON), 14 from the High Rain Wheat Yield Trial (HRWYT), 34 from the High Rainfall Wheat Screening Nursery (HRWSN), 5 from an adaptation trial, 6 from the National Variety Trials (NVT), and the remaining 3 genotypes were from a preliminary variety trial (PVT).

Multi-Environment Trials

Field evaluations were carried out under natural STB infestation during the 2015 and 2016 main cropping seasons across three STB hotspots: the Holetta Agricultural Research Center (HARC) (9° 3'N/38° 30'E), Bekoji Agricultural Research Subcenter (7° 32'N/39° 15'E), and Kulumsa Agricultural Research Center (KARC) (8° 02'N/ 39° 15'E). The experimental design was an alpha lattice with two replications, six incomplete blocks, and 30 entries per sub-block per replication. The trial was sown by hand, with each entry planted in four rows of a 2.5-m length, 20-cm spacing between rows, and 40 cm between entries. The susceptible cv. "Lakech" was planted as a spreader row along the length of the blocks to create adequate disease pressure. The spaces between the blocks and replications were 1.5 m long. A seeding rate of 150 kg ha⁻¹ and fertilizer rates of 100 and 75 kg ha⁻¹ of N and P₂O₅, respectively, were used in all the experiments. Weeding was performed by hand three times each season.

STB Evaluation

Septoria tritici blotch disease severity (SDS) was estimated visually plot-wise by considering the percentage of necrotic leaf area (NLA) on the four uppermost infected leaves of 10–20 plants (Eyal and Levy, 1987) at three growth stages, namely, heading (SDH), medium milk (SDMM), and at maturity (SDM), using a double-digit 00–99 scoring scale (Eyal and Levy, 1987). The first digit (0–9) represented blotch development in terms of plant height (for instance, 5 if the disease reached the middle (50%) of the plant height, 8 for reaching the flag leaf, and 9 for reaching the spike), while the second digit stood for the disease severity as a percentage but in terms of 0–9 (1 = 10%, 2 = 20%, and 9 = 90 %). For each stage, Septoria disease severity percentage (SDS%) was computed from the 00–99 score using the following formula as described by Sharma and Duveiller (2007):

$$SDS = [(D1/9) (D2/9)] \times 100$$

where D1 and D2 are the first and the second digits of the double-digit scores, respectively. The SDS values range from 0 to 100, where 0 indicates complete resistance and 100 indicates complete susceptibility (Kidane et al., 2017).

In addition, the Septoria progress coefficient (SPC) developed by Eyal and Ziv (1974) was computed to indicate the position of pycnidia relative to plant height according to the following equation:

$$SPC = (SDH/PH)$$

where SDH (Septoria disease height) is the maximum height (cm) above ground at which the pycnidia of the pathogen could be found on the plant at the maturity stage and PH is the average height of the genotype from the ground to the tip of its awn. The SPC coefficient indicates the position of pycnidia relative to plant height, regardless of pycnidial coverage, and allows for the comparison of the infection placements on cultivars with different plant statures. Furthermore, SPC values ranged from 0 to 1, where SPC = 0 means that there was no disease, while SPC = 1 means that pycnidia were produced at the top of the plant (Eyal and Levy, 1987).

Other Agronomic Data Scoring

The phenotypic data that were recorded were heading date (HD, days to 50% heading), flowering date (FD = days to 50% flowering), grain-filling duration (GFD), maturity date (MD = days to 90% maturity), grain yield, hectoliter weight (HLW = kilograms per 100 liters of wheat), thousand-kernel weight (TKW = weight of 1,000 kernels, in grams), plant height (PH), number of spikelets per spike (SPS), number of kernels per spikelet (NKPS), and number of kernels per spike (NKS). These yield data were taken from the four rows of each plot and converted to kilograms per hectare (kg ha⁻¹) at 12.5% moisture content using plot size as a factor. Plant height measurement was also carried out at physiological maturity from five randomly selected and tagged plants from the middle rows of each entry.

DNA Extraction and Genotyping by Sequencing

The wheat plants of the association panel were grown at the National Agricultural Biotechnology Research Center, Holetta under greenhouse conditions. The 2-week-old leaf samples were then collected into 96 deep-well sample collection plates, oven-dried overnight at 50°C, and sent to Integrated Genotyping Service and Support (IGSS) located at the Biosciences Eastern and Central Africa (BecA-ILRI) Hub in Nairobi, Kenya for high-density genotyping by Diversity Arrays Technology sequencing (DARtSeq™ technology). Furthermore, DNA extraction was carried out using the Nucleomag Plant Genomic DNA extraction kit (Macherey-Nagel GmbH & Co. KG, Duren, Germany). Afterward, extracted DNA quality and quantity were checked on a Thermo Scientific™ NanoDrop™ 2000 Spectrophotometers (Thermo Scientific™, USA) and on 0.8% agarose gels. As a result, the extracted genomic DNA concentration ranged from 50 to 100 ng/μl. Whole-genome profiling was also carried out using the genotyping-by-sequencing (GBS) platform as described by Elshire et al. (2011). This method involved library construction following the DARTSeq complexity reduction method *via* the digestion of genomic DNA using *ApeKI* [a type II restriction endonuclease that recognizes a degenerate 5-bp sequence (GCWGC, where W is A or T)] and the ligation of barcoded adapters, which was also followed by the PCR amplification of adapter-ligated fragments. The libraries were then sequenced using single-read sequencing runs for 77 bases. The next-generation sequencing of the GBS library was also carried out using an Illumina HiSeq 2500 lane (Illumina, San Diego, CA, United States) following the protocol of the manufacturer.

Quality Control and SNP Calling

The technical quality of the sequencing was checked using a Sequencing Analysis Viewer. DArTSeq markers were scored using the *DArTsoft14* software implemented in the KDCompute plug-in system developed by Diversity Arrays Technology (2017) (<http://www.kddart.org/kdcompute.html>) based on their alignment with the reference genome of the Chinese Spring Wheat RefSeq v1.0 [International Wheat Genome Sequencing Consortium (IWGSC), 2018], which was obtained from the International Wheat Genome Sequencing Consortium database (<https://urgi.versailles.inra.fr/download/iwgs/>) at a minimum base identity of 90% and e-value of $5e-10$. Two types of markers were scored, namely, SilicoDArT markers and SNP markers, which were both scored in a binary fashion (1/0), indicating the presence or absence of a marker in the genomic representation of each sample as described by Akbari et al. (2006). Marker quality was also maintained by removing monomorphic markers and those with lower call rates ($>30\%$ missing) and MAFs (minor allele frequencies) $<5\%$ using the ArTSoft14 software.

Statistical Data Analysis

Phenotypic Data Analysis

We conducted an ANOVA for each location in each year using the SAS software version 9.2 (SAS Institute Inc., 2008) by considering genotype and the block as fixed and random factors, respectively. In an individual environment, the observed phenotypic response of the i th genotype in the j th replication and l th sub-block was computed using the following model:

$$y_{ijl} = \mu + g_i + \gamma_j + bl_{(j)} + \epsilon_{ij} \quad (1)$$

where y_{ijl} = the observed phenotype, μ = the grand mean, g_i = fixed effect of the i th genotype, γ_j = effect of the j th replication, $bl_{(j)}$ = random effect of the l th block nested within the j th replication, and ϵ_{ijl} = random error term.

The ANOVA of all seasons and locations was executed by considering genotype as a fixed effect and the block, location, and year as random effects according to the following model:

$$Y_{ijklm} = \mu + g_m + \gamma_{ijk} + \gamma_{ij} + e_j + b_{ijkl} + (gy)_{im} + (ge)_{jm} + (ye)_{ij} + (yeg)_{ijm} + \epsilon_{ijklm}$$

where Y_{ijklm} = observed response of genotype m , replication k of block l of location j and year i ; μ = grand mean; g_m = fixed effect of genotype m ; r_{ijk} = effect of replication k in location j and year i ; γ_{ij} = random effect of year i at location j that is \sim normally and independently distributed (NID) $(0, \delta_y^2)$; e_j = random effect of location j that is \sim NID $(0, \delta_e^2)$; b_{ijkl} = random effect of block l nested with replication k in location j and year i that is \sim NID $(0, \delta_b^2)$; $(gy)_{im}$ = random effect of the interaction between genotype m and year i that is \sim NID $(0, \delta_{gy}^2)$; $(ge)_{jm}$ = random effect of the interaction between genotype m and location j that is \sim NID $(0, \delta_{ge}^2)$; $(ye)_{ij}$ = random effect of the interaction between year i and location j that is \sim NID $(0, \delta_{ye}^2)$; $(yeg)_{ijm}$ = random effect of the three-way interaction of genotype m in location j and year i that is \sim NID $(0, \delta_{yeg}^2)$; ϵ_{ijklm} = random residual effect that is \sim NID $(0, \delta_e^2)$.

The variance components were also computed. The broad-sense heritability (H^2) within an environment was estimated for the traits from an ANOVA using the following formula:

$$H^2 = (\delta^2_g) / (\delta^2_g + \delta^2_e / r)$$

The broad-sense heritabilities across the environments were also estimated by the formula:

$$H^2 = (\delta^2_g) / (\delta^2_g + \delta^2_{gy} / y + \delta^2_{ge} / l + \delta^2_{gye} / yl + \delta^2_e / ylr)$$

where δ^2_g is the genotypic variance, σ^2_{gy} is the genotype-by-year interaction variance, σ^2_{ge} is the genotype-by-location interaction variance, σ^2_{gye} is the genotype-by-year-by-location interaction variance, δ^2_e is the location variance, and l , r , and y represent the numbers of locations, replicates, and years, respectively. The percentage of heritability was categorized as low ($<30\%$), moderate ($30-60\%$), or high ($\geq 60\%$) as described by Robinson et al. (1949). The relationship between agronomic traits was also determined by Pearson's correlation using the SAS software version 9.2 (SAS Institute Inc., 2008).

Population Structure Analysis

A population stratification of the association panel was visualized by principal component analysis (PCA) using the KDCompute plug in system version 1.0.1 (<https://kdcompute.seqart.net/kdcompute/plugins>). Population admixture patterns were also determined using a Bayesian model-based clustering algorithm implemented in the STRUCTURE software v.2.3.4 (Pritchard et al., 2000). The STRUCTURE program was run with the admixture model, correlated allele frequencies, a burn-in period of 10,000, and 50,000 Markov Chain Monte Carlo (MCMC) replications after a burn in for hypothetical subpopulations K from 1 to 10 with 10 iterations. The optimum K value was predicted based on a study by Evanno et al. (2005) using STRUCTURE HARVESTER ver. 0.6.92 (Earl and von Holdt, 2012). A bar plot for the optimum K was determined using Clumpak beta version (Kopelman et al., 2015).

Genome-Wide Association Study

The association mapping of phenotypic traits with genome-wide scanned SNPs was conducted using the Genome Association and Prediction Integrated Tools (GAPIT) package (Lipka et al., 2012) in the R software (R Core Team, 2013). This GWAS was carried out for four Septoria disease traits, namely, SDSH, SDSMM, SDSM, and SPC, and some important agronomic traits such as the days to 50% heading (DH), days to 50% flowering (DF), grain filling duration (GFD), days to 90% maturity (DM), grain yield, thousand-kernel weight (TKW), and plant height (PH) in each individual environment; the study design also used the means across all environments [the best linear unbiased estimate (BLUE) values]. The analysis involved a total of 7,776 robust SNPs with a call rate of $>70\%$ and MAF of $>5\%$. Missing SNP data were imputed using optimal impute ver. 1.0.0 in the KDcompute_plugin system based on the KNN imputation method. The marker distribution on each chromosome was determined using LD measure in R^2 ver.0.2.2 of the KDcompute_plugin. Pairwise LD measures (r^2 and P -value) between markers on each chromosome were also computed using

TASSEL Ver. 5 (Bradbury et al., 2007). A genome-wide LD decay scatter plot was then produced by plotting the r^2 values against physical distance (bp) using the GAPIT software. Finally, $r^2 = 0.2$ was considered as a cutoff point for no LD between pairs of markers.

The GWAS was conducted using the fixed and random model circulating probability unification (FarmCPU) algorithm (Liu et al., 2016) implemented in the GAPIT R package (2.0) (Tang et al., 2016). The algorithm uses both fixed-effect and random-effect models iteratively to control spurious marker-trait associations due to population structure and family relatedness (Lipka et al., 2012). Furthermore, a kinship (K) matrix was computed using the method of VanRaden (2008). Principal components describing the population stratification were computed using R/GAPIT and iteratively added to the fixed part of the model. Quantile–quantile (Q–Q) plots generated from $-\log_{10} p$ -values were assessed visually to determine how well the model accounted for population structure and family relatedness among the study samples. Statistically significant marker-trait associations were declared using a false discovery rate (FDR)-adjusted $p < 0.05$ as implemented in GAPIT. Furthermore, the Bonferroni correction rate at a significance threshold of $p < 0.15$ or $-\log_{10} (p\text{-values}) = 4.71$ was also included in the analysis for comparison. Both the Q–Q and Manhattan plots were visualized using the R package qqman (Turner, 2014). The high-confidence candidate genes within the identified resistance-associated regions were also extracted from the recently released IWGSC RefSeq Annotation v2.1 available on the URGI Seq repository (<https://wheat-urgi.versailles.inrae.fr/Seq-Repository/Annotations>).

RESULTS

Phenotypic Data Analysis

Adult Plant Responses to STB and Broad-Sense Heritability

The genotype effect was significant ($p < 0.0001$) for STB resistance at all the growth stages in all the test environments. Genotypic variance (σ^2_g) was the major contributor to STB resistance variability among the tested wheat genotypes. The Septoria disease severity traits also showed pseudo-normal distributions (Figure 1), indicating the quantitative nature of STB resistance in the tested wheat genotypes (Kidane et al., 2017). The analysis revealed that the STB infestation showed seasonal fluctuations, but that was still higher during the 2015 growing season (Table 1). Moreover, the disease severity showed an increasing trend from heading to the maturity stage. In each environment, mean SDS values at the heading and mid-maturity stages ranged from 18.2 to 31.2% and 21.7 to 37.6%, respectively, while the highest severity values were registered at Holetta in 2015. The mean disease severity at maturity and its vertical progress varied from 30 to 50.8% and 0.41 to 0.69, respectively, while they were the highest at Bekoji in the 2015 growing season. The lowest Septoria severity was recorded at Kulumsa in 2016. The broad-sense heritability for Septoria resistance in each

environment ranged from moderate ($H^2 = 0.58$) to high ($H^2 = 0.99$) (Table 2).

The combined ANOVA revealed that the effect of genotype, year, location, and their two- (genotype \times year, genotype \times location, and year \times location) and three-way interactions (genotype \times year \times location) were significant for SDS traits except for the Septoria progress coefficient, where the effect of year was not significant (Table 3). The analysis of pooled data revealed that genotypic variance (σ^2_g) was the highest for all the SDS parameters except SDSMM, where environmental effect was the highest (Table 4). The broad-sense heritabilities of Septoria resistance traits showed that they were highly heritable ($H^2 = 72\text{--}81\%$) (Table 4) (Robinson et al., 1949). The phenotypic and genotypic coefficients of variation for SDS traits ranged from 32.4 (SPC) to 68.3% (SDSH) and 22.5 (SPC) to 41.4% (SDSM), respectively. At 5% selection intensity, the genetic advance for SDS traits ranged from 0.31 (SPC) to 35.23 (SDSM), while the magnitude of the expected genetic gains as a percent of the mean varied from 53.69% (SPC) to 102.38% for SDSH (Table 4).

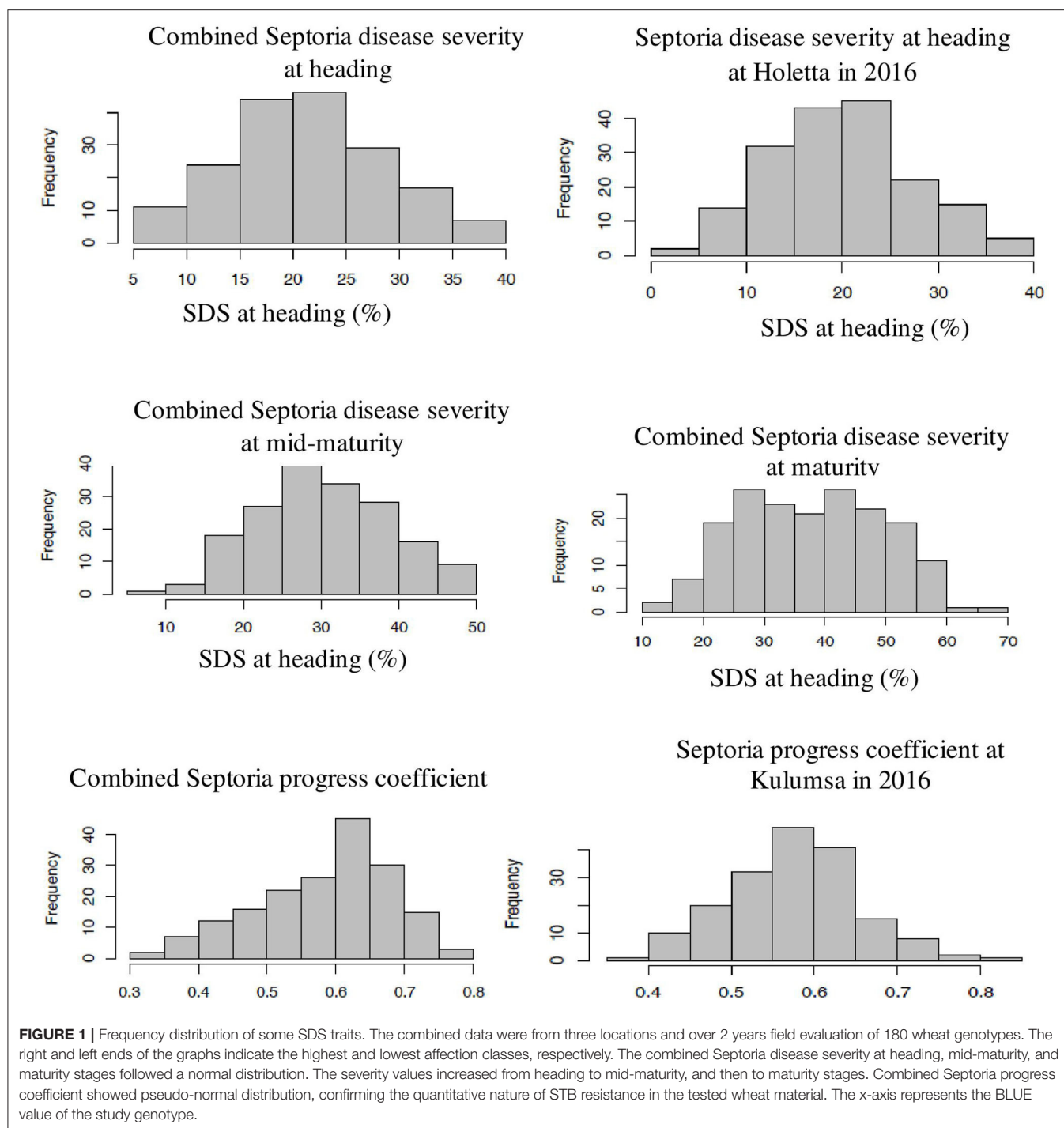
Over all the environments, the average SDS of the individual wheat genotypes ranged from 5.3 to 39.8% at heading, 8.2 to 48.5% at mid-maturity, and 10.6 to 65.3% at maturity (Supplementary Table 2). The average SPC of the individual environments ranged from 0.37 to 0.79. The most resistant genotype at all the growth stages was G174, while G104 (39.8%), G76 (48.5%), and G127 (65.3%) were the most susceptible genotypes at the heading, mid-maturity, and maturity stages, respectively (Supplementary Table 2). A comparative severity analysis with the standard check King-bird (G40) and the mean performance of the released varieties also confirmed the presence of superior STB-resistant genotypes among the tested materials. Of the 180 tested genotypes, 56 (31%) at heading, 75 (42%) at mid-maturity, and 105 (59%) at maturity had numerically superior STB resistance compared with King-bird (Supplementary Table 2).

The top 5% best genotypes at maturity had 47.6–71% greater resistances than King-bird and 11.9–74.4 % greater resistances compared with the mean performances of the released varieties (Table 5).

Pearson's correlation analysis of the means over all environments revealed that STB resistance traits were significantly negatively associated with important agronomic traits. Except for SPC, all the SDS traits showed non-significant and negligible negative correlations ($r < -0.3$) with plant height. Disease traits also showed little to negative associations with HD, FD, GFD, NKPS, and NKS. However, a significant weak negative association (-0.25 to -0.48) was observed between SDS traits and MD, grain yield, HLW, and TKW (Table 6).

SNP Statistics

The Illumina HiSeq 2500 (Illumina, San Diego, CA, United States) sequencing failed to generate SNP data for two genotypes (9 and 95); hence, a total of 178 bread wheat genotypes were successfully DArTSeq genotyped. Initially, a total of 35,672 SNPs were discovered, of which 31,052 (87%) were mapped to known chromosomal positions on the reference used and 828 (2%) of the sequences were mapped to unknown



chromosomes in the reference. In contrast, approximately 11% (3,792) of the SNPs did not align to any of the chromosomes of the wheat reference genome. Furthermore, the discovered DArTSeq SNPs were not evenly distributed among the sub-genomes, with the A, B, and D genomes accounting for 10,317, 10,979, and 9,756 SNPs, respectively (**Supplementary Figure 1**). Among the 21 wheat chromosomes, the highest (2,065) and the lowest (833) numbers of SNPs were mapped to chromosomes 7D

and 4D, respectively (**Supplementary Figure 1**), and on average, each chromosome harbored about 1,479 SNPs. Maintaining SNPs with higher call rate (>70%) and MAF >0.05 resulted in 7,776 SNP markers, among which 87.3% had a known chromosome position in the wheat reference genome. Among the filtered SNPs, 2,410 were distributed on the A genome, 2,872 were distributed on the B genome, and 1,506 were distributed on the D genome. The remaining 988 SNPs were assigned to a

TABLE 1 | Descriptive statistics of SDS values in bread wheat genotypes evaluated in three locations in Ethiopia during the 2015 and 2016 growing seasons.

| Trait | Environment | Mean | Range | SD | Pr > F | Trait | Environment | Mean | Range | SD | Pr > F |
|-------|-------------------|-------|-------|------|--------|-------|-------------------|-------|-------|-------|--------|
| SDSH | E1 (HARC2015) | 31.23 | 72.84 | 8.88 | *** | SDSM | E1 (HARC 2015) | 46.9 | 95.68 | 24.07 | *** |
| | E2 (Bekoji 2015) | 24.48 | 63.58 | 6.8 | *** | | E2 (Bekoji 2015) | 50.77 | 77.16 | 19.05 | *** |
| | E3 (Kulumsa 2015) | 23.47 | 40.74 | 8.00 | *** | | E3 (Kulumsa 2015) | 32.38 | 70.99 | 9.44 | *** |
| | E4 (HARC 2016) | 20.09 | 39.81 | 4.88 | *** | | E4 (HARC 2016) | 33.07 | 65.44 | 1.8 | *** |
| | E5 (Bekoji 2016) | 19.36 | 44.71 | 9.08 | *** | | E5 (Bekoji 2016) | 34.28 | 55.56 | 10.41 | *** |
| | E6 (Kulumsa 2016) | 18.16 | 38.89 | 6.87 | *** | | E6 (Kulumsa 2016) | 30.04 | 46.94 | 9.61 | *** |
| SDSMM | E1 (HARC 2015) | 37.6 | 82.72 | 22.7 | *** | SPC | E1 (HARC 2015) | 0.65 | 0.87 | 0.2 | *** |
| | E2 (Bekoji 2015) | 34.56 | 74.07 | 17.2 | *** | | E2 (Bekoji 2015) | 0.69 | 0.74 | 0.12 | *** |
| | E3 (Kulumsa 2015) | 32.13 | 49.51 | 8.2 | *** | | E3 (Kulumsa 2015) | 0.41 | 0.60 | 0.09 | *** |
| | E4 (HARC 2016) | 27.13 | 58.95 | 5.29 | *** | | E4 (HARC 2016) | 0.6 | 0.62 | 0.09 | *** |
| | E5 (Bekoji 2016) | 25.6 | 47.03 | 9.51 | *** | | E5 (Bekoji 2016) | 0.58 | 0.61 | 0.12 | *** |
| | E6 (Kulumsa 2016) | 21.73 | 39.51 | 8.35 | *** | | E6 (Kulumsa 2016) | 0.58 | 0.46 | 0.08 | *** |

SDSH, Septoria disease severity at heading; SDSMM, Septoria disease severity at mid- maturity; SDSM, Septoria disease severity at maturity; SPC, Septoria progress coefficient; SD, standard deviation; ***, very highly significant ($p \leq 0.0001$); "sn", non-significant at the α , 5% significance level.

TABLE 2 | Genotypic variance (σ^2g) and heritability in the broad sense (H^2) for phenotypic traits in 180 bread wheat genotypes in six different environments in Ethiopia.

| Trait | Holetta-2015 (E1) | | Bekoji-2015 (E2) | | Kulumsa-2015 (E3) | | Holetta-2016 (E4) | | Bekoji-2016 (E5) | | Kulumsa-2016 (E6) | |
|-------|-------------------|-------|------------------|-------|-------------------|-------|-------------------|-------|------------------|-------|-------------------|-------|
| | σ^2g | H^2 | σ^2g | H^2 | σ^2g | H^2 | σ^2g | H^2 | σ^2g | H^2 | σ^2g | H^2 |
| SDSH | 344.51*** | 0.90 | 212.66*** | 0.90 | 39.14*** | 0.61 | 42.56*** | 0.78 | 79.41*** | 0.96 | 44.67*** | 0.95 |
| SDSMM | 485.36*** | 0.94 | 255.28*** | 0.86 | 38.7*** | 0.58 | 92.23*** | 0.90 | 88.63*** | 0.98 | 66.48*** | 0.95 |
| SDSM | 516.40*** | 0.89 | 275.44*** | 0.76 | 331.16*** | 0.88 | 127.99*** | 0.99 | 105.24*** | 0.97 | 86.59*** | 0.94 |
| SPC | 0.04*** | 0.88 | 0.01*** | 0.70 | 0.02*** | 0.87 | 0.01*** | 0.62 | 0.01*** | 0.83 | 0.00*** | 0.67 |

σ^2g , genotypic variance; H^2 , heritability in the broad sense (H^2); ***, very highly significant ($p \leq 0.0001$); "sn" = non-significant at the α , 5% significance level.

TABLE 3 | Combined analysis of variance for Septoria disease severity traits across three locations in Ethiopia over years 2015 and 2016.

| Source of variation | DF | Mean squares | | | |
|---------------------------|-----|--------------|----------------------|-------------|--------------------|
| | | SDSH | SDSMM | SDSM | SPC |
| Genotype | 179 | 699.53*** | 874.84*** | 1650.29*** | 0.12*** |
| Replication | 1 | 662.48* | 35.41 ^{ns} | 5498.56*** | 0.87*** |
| Year | 1 | 29489.27*** | 44295.09*** | 64059.86*** | 0.01 ^{ns} |
| Location | 2 | 5257.27*** | 2734.03 *** | 24849.55*** | 4.57*** |
| Incomplete block (nested) | 5 | 1468.79*** | 2167.8*** | 2955.71*** | 0.31*** |
| Genotype*year | 179 | 310.41*** | 392.85*** | 1008.41*** | 0.06*** |
| Genotype*location | 358 | 207.74*** | 288.83*** | 172.91*** | 0.02*** |
| Year*location | 2 | 1987.73*** | 358.46 ^{ns} | 10452.82*** | 3.78*** |
| Genotype*Year*Location | 358 | 148.97*** | 213.74*** | 152.76* | 0.02*** |

SDSH, Septoria disease severity at heading; SDSMM, Septoria disease severity at mid-maturity stage; SDSM, Septoria disease severity at maturity; SPC, Septoria progress coefficient; ***, very highly significant at $p < 0.0001$; *, significant at $p < 0.05$; ns, non-significant at the $p = 0.05$ significance level; DFs, degrees of freedom.

hypothetical chromosome "0" for the sake of analysis. Hence, 7,776 SNPs were used in downstream analyses, which included principal component analysis (PCA), population clustering, population structure, LD, and GWAS.

Population Structure Analysis

The STRUCTURE analysis indicated two sub-populations in the association panel (Figure 2A), where ~43% (76) of the genotypes

were assigned to cluster one and 57% (101) were assigned to cluster two. Additionally, the Clumpak result detected a greater degree of genetic admixture between the two sub-populations (Figure 2B), where all the individual genotypes shared alleles inherited from both subgroups (Figure 2C), thus confirming the presence of close relationships among the study materials. Furthermore, the PCA results also suggested the presence of two sub-populations (Figure 3). The first two principal components

TABLE 4 | Variance component estimates for SDS, H^2 (broad sense), genotypic coefficient of variance (GCV), phenotypic coefficient of variance (PCV), genetic advance (GA), and genetic advance as percent of the mean (GAM) based on pooled data from the six environments.

| Trait | σ^2g | σ^2gy | σ^2gl | σ^2gyl | σ^2e | H^2 | GCV | PCV | GA | GAM |
|-------|-------------|--------------|--------------|---------------|-------------|-------|-------|-------|-------|--------|
| SDSH | 81.97*** | 26.91*** | 14.7*** | 34.63*** | 79.71*** | 0.73 | 40.07 | 68.26 | 23.14 | 102.38 |
| SDSMM | 97.67*** | 29.86*** | 18.78*** | 40.46*** | 132.82*** | 0.72 | 32.62 | 59.01 | 26.36 | 86.99 |
| SDSM | 246.24*** | 142.61*** | 5.04*** | 15.32*** | 122.13*** | 0.75 | 41.41 | 60.83 | 35.23 | 93.09 |
| SPC | 0.02*** | 0.01*** | -0.01*** | 0.01*** | 0.01*** | 0.81 | 22.51 | 32.37 | 0.31 | 53.69 |

σ^2g , genotypic variance estimate; σ^2gl , genotype \times year interaction variance estimate; σ^2gyl , genotype \times year \times location interactions variance estimate; σ^2e , residual variance estimate; ***, very highly significant at $p < 0.0001$; *, significant at $p < 0.05$; ns, non-significant at the $p = 0.05$ significance level; SDSH, Septoria disease severity at heading; SDSMM, Septoria disease severity at mid maturity stage; SDSM, Septoria disease severity at maturity; SPC, Septoria progress coefficient.

TABLE 5 | Comparison of the mean performances of 5% of the genotypes selected for Septoria tritici blotch (STB) resistance with King-bird, a recently released variety, and with the mean performances of 13 released varieties.

| Genotypes | Mean of selected genotypes | Comparative advantage for STB resistance (% over) | | Genotypes | Mean of selected genotypes | Comparative advantage for STB resistance (% over) | |
|---|----------------------------|---|-------|---|----------------------------|---|-------|
| | | King-bird | MRV* | | | King-bird | MRV* |
| Septoria disease severity at heading (%) | | | | Septoria disease severity at maturity (%) | | | |
| G174 | 5.25 | 70.95 | 74.4 | G174 | 10.6 | 74.25 | 71.83 |
| G153 | 5.82 | 67.81 | 71.63 | G144 | 14.21 | 65.49 | 62.24 |
| G144 | 5.87 | 67.53 | 71.38 | G3 | 15.23 | 63 | 59.51 |
| G150 | 6.07 | 66.39 | 70.38 | G153 | 15.95 | 61.25 | 57.6 |
| G151 | 8.45 | 53.22 | 58.78 | G156 | 18.73 | 54.5 | 50.21 |
| G141 | 8.9 | 50.72 | 56.57 | G133 | 18.83 | 54.25 | 49.94 |
| G133 | 8.9 | 50.72 | 56.57 | G151 | 19.04 | 53.75 | 49.4 |
| G3 | 9.32 | 48.44 | 54.56 | G155 | 19.45 | 52.75 | 48.3 |
| G156 | 9.47 | 47.58 | 53.81 | G97 | 19.66 | 52.25 | 47.75 |
| King-bird | 18.06 | – | 11.88 | King-bird | 41.15 | – | –9.43 |
| MRV* | 20.49 | 13.48 | – | MRV* | 37.61 | 8.62 | – |
| Septoria disease severity at mid-maturity (%) | | | | Septoria progress coefficient | | | |
| G174 | 8.18 | 70.4 | 72.74 | G174 | 0.31 | 50.44 | 45.59 |
| G153 | 10.55 | 61.83 | 64.86 | G144 | 0.35 | 44.51 | 39.09 |
| G144 | 14.15 | 48.79 | 52.85 | G133 | 0.37 | 40.82 | 35.03 |
| G3 | 14.46 | 47.68 | 51.83 | G3 | 0.37 | 40.31 | 34.48 |
| G151 | 15.28 | 44.7 | 49.08 | G155 | 0.37 | 40.3 | 34.46 |
| G155 | 15.75 | 43.02 | 47.54 | G151 | 0.38 | 38.97 | 33.01 |
| G92 | 16.31 | 40.97 | 45.65 | G154 | 0.39 | 38.03 | 31.97 |
| G150 | 16.36 | 40.79 | 45.48 | G97 | 0.39 | 37.76 | 31.68 |
| G81 | 17.03 | 38.37 | 43.25 | G47 | 0.4 | 36.88 | 30.71 |
| King-bird | 27.63 | – | 7.93 | King-bird | 0.62 | – | 9.78 |
| MRV* | 30.01 | –8.61 | – | MRV* | 0.57 | 8.91 | – |

*MRV, mean of 13 selected released varieties. Negative values for comparative advantage indicate less STB resistance (inferior performance) of the genotype.

explained 65% ($PC1 = 50\%$ and $PC2 = 15\%$) of the total variance contained in the data (Figure 3). With this, a scree plot, which was used to display the proportion of variation captured by each of the 10 principal components, also showed that the first two principal components ($PC1$ and $PC2$) explained

the highest proportion of the total variation in the panel (Figure 4A). Figure 4B represents the 3D plots of the first three principal components to depict the samples' relationship in space, the analysis also confirmed the presence of kinship in the association panel (Figure 4C), suggesting the importance

TABLE 6 | Correlation analyses among Septoria resistance traits and some agronomic traits in 180 bread wheat genotypes based on the pooled data from 2 years of field trials in Ethiopia.

| | HD | FD | MD | GFD | Yield | HLW | TKW | PH | NKPS | NKS |
|-------|---------------------|---------------------|----------|---------|----------|----------|----------|---------------------|---------------------|---------------------|
| SDSH | -0.04 ^{ns} | -0.05 ^{ns} | -0.19* | -0.21* | -0.43*** | -0.32*** | -0.42*** | -0.1 ^{ns} | -0.19* | -0.19* |
| SDSMM | -0.1 ^{ns} | -0.1 ^{ns} | -0.25* | -0.24** | -0.48*** | -0.33*** | -0.51*** | -0.08 ^{ns} | -0.25** | -0.21* |
| SDSM | -0.21** | -0.22* | -0.35*** | -0.26** | -0.47*** | -0.32*** | -0.44*** | -0.06 ^{ns} | -0.16* | -0.14 ^{ns} |
| SPC | -0.15* | -0.14 ^{ns} | -0.29*** | -0.22* | -0.38*** | -0.25** | -0.44*** | -0.21* | -0.12 ^{ns} | -0.1 ^{ns} |

HD, heading date; FD, flowering date; MD, maturity date; GFD, grain-filling duration; HLW, hectoliter weight; TKW, thousand-kernel weight; PH, plant height; NKPS, number of kernels per spikelet; NKS, number of kernels per spike; SDSH, Septoria disease severity at heading; SDSMM, Septoria disease severity at mid maturity stage; SDSM, Septoria disease severity at maturity; SPC, Septoria progress coefficient.

***, very highly significant ($p < 0.0001$); ** = highly significant; *, significant; ns, non-significant at the α , 5% significance level; (-), negative correlation. The magnitude of the correlation coefficient indicates the strength of the correlation.

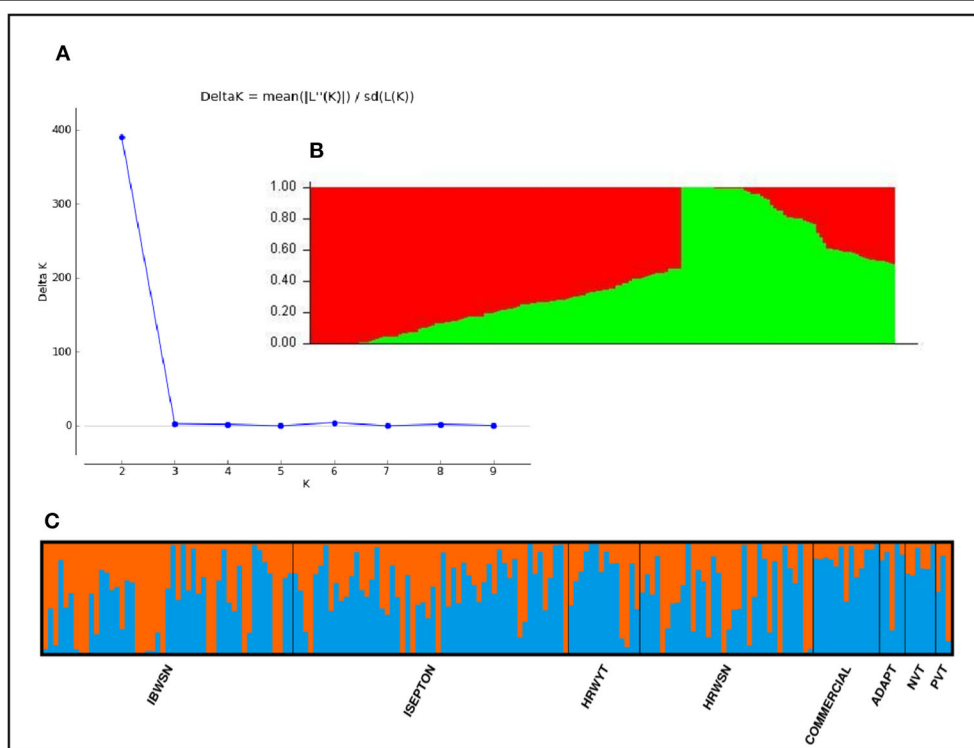


FIGURE 2 | Population structure of the 178 bread wheat genotypes representing eight populations. **(A)** Best delta K value estimated using the method of Evanno et al. (2005), and the pick at $k = 2$ indicates the number of subpopulations in our collection, **(B)** Population structure plot and SP1 and SP2 represents subpopulations 1 and 2, respectively, **(C)** Estimated population structure for $K = 2$ according to the breeding materials. The different (blue and orange) colors represent genetic groups or sub-populations designated by Structure Harvester: the x-axis represents individual samples and y-axis represents the proportion of ancestry to each cluster. Population abbreviations are: IBWSN, International Bread Wheat Screening Nursery; ISEPTON, International Septoria Observation Nursery; HRWYT, High Rain Wheat Yield Trial; HRWSN, High Rain Wheat Screening Nursery; ADAPT, Adaptation trial; NVT, National Verification Trial, and PVT, Preliminary Verification.

of using a powerful statistical GWAS model that accounts for the population structure and familial relatedness in the association study.

Linkage Disequilibrium (LD) Analysis

The linkage disequilibrium of alleles at different loci varied considerably across each chromosome and among chromosomes and sub-genomes (Table 7). There was a total of 338,125 marker pairs with average LD values of $r^2 = 0.11$, with 97,723 (27.6%) pairs showing significant linkage at $p \leq 0.01$ (Table 7). In

particular, the B genome harbored the highest (143,600 or 42.5%) number of marker pairs, followed by the A genome with 119,225 (35.5%) of the marker pairs (Table 7). In contrast, the D genome harbored the lowest number (75,300 or 22.3%) of the marker pairs. Relatively, however, the SNPs on the B genome showed the strongest LD, with a mean value of $r^2 = 0.1187$. Over all the chromosomes, the LD between SNPs declined to $r^2 = 0.2$ within a physical distance of 31.44 Mbp; this ranged from 2.26 to 105.6 Mbp by chromosome. The weakest and strongest LD values were observed between the marker pairs on chromosomes 4D ($r^2 =$

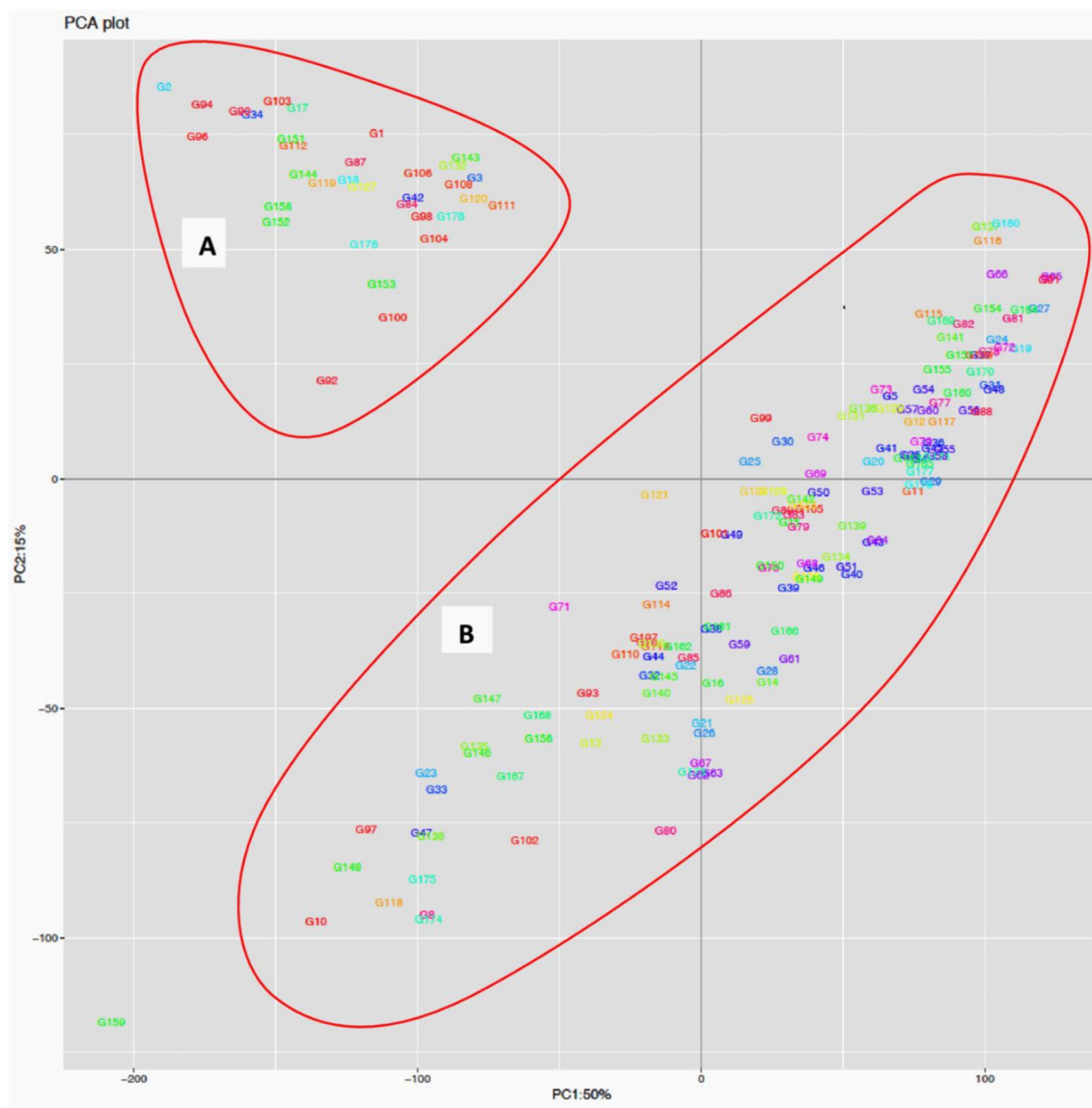


FIGURE 3 | Population structure analysis of the 178 bread wheat genotypes based on principal component analysis clustering as revealed by the first two principal components. Samples coded with the same color belong to same population. Cluster **(A)** composed of 33 (18.54%) genotypes while Cluster **(B)** possessed 145 (81.46%) of the genotypes.

0.0251) and 2D ($r^2 = 0.211$), respectively (Table 7). The physical distance (bp) at which the LD decayed to the critical r^2 (0.2) value was used to determine the confidence interval for declaring the distinct QTL for each chromosome. Significant SNP markers from the same chromosome were also assigned to the same QTL if the distance between the significant markers was less than the critical physical distance.

Genome-Wide Association Study STB Resistance

The GWAS identified 53 SNPs that were significantly (FDR < 0.05) associated with STB resistance at any growth stage. The report, however, only included the marker-trait

associations (MTAs) that surpassed a Bonferroni-correction significance threshold of 0.15. **Supplementary Table 3** presents the complete output of the GWAS results for STB resistance at the heading, mid-maturity, and maturity stages and for the Septoria progress coefficient. This table also reports allele identity, marker position, MAF, p -values, FDR-corrected q values, and the additive effects of the identified MTAs. The Q-Q plots demonstrating how well the used GWAS model accounted for population structure and kinship for STB resistance analysis are also presented in **Supplementary Figure 2**.

Among the 53 identified MTAs, 3 did not have chromosomal positions on the bread-wheat physical map

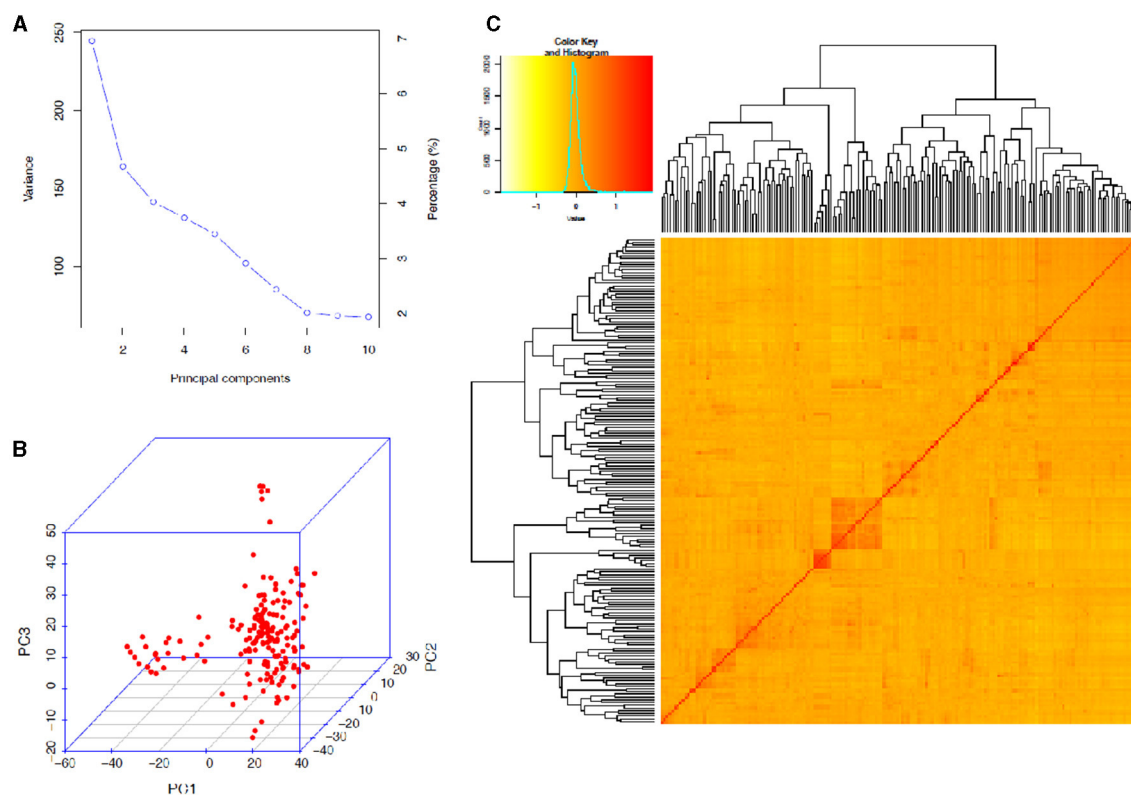


FIGURE 4 | Principal component and familiar relatedness analyses of 178 bread wheat genotypes based on 7,776 genome-wide scanned, high-quality SNPs. **(A)** A screen plot displaying the first 10 principal components with their corresponding fraction of variation explained, **(B)** 3D plots of the first three principal components to depict the samples' relationship in space, and **(C)** Heat map showing the kinship analysis. The kinship values showed a normal distribution (turquoise curve), and orange and red colors represent weak and high kinship relations in the panel, respectively. The resulted clustering tree is indicated outside of the matrix.

(**Supplementary Table 4**). Ten (18.9%) of the MTAs conferred STB resistance at heading, 4 (7.6%) were effective at mid-maturity, 4 (7.6%) were effective at the maturity stage, and 35 (66.9%) of the MTAs were associated with a resistance to disease development as plant height increased (**Supplementary Table 4**). The percentage of phenotypic variance explained by the markers varied considerably, from 2.7% for the SDS measured at maturity at Bekoji in 2016 to 13.2% for the severity data measured at the mid-maturity stage at the same location in 2015 (**Supplementary Table 4**). The proportion of phenotypic variation (R^2) explained by SDS MTAs at heading ranged from 2.9% for the allele 1195254[F|0-31:C>T-31:C>T on chromosome 3A to 11.1% for 1087857[F|0-41:T>C-41:T>C on chromosome 7D (**Supplementary Table 4**). Likewise, the R^2 for MTAs for SDS at mid-maturity, maturity, and SPC ranged from 8.6 to 13, 2.7 to 2.7, and 6 to 10.8%, respectively (**Supplementary Table 4**).

The combined measure of SDS at the heading, mid-maturity, and maturity stages did not provide any significant associations at the used threshold. However, the combined measure of SPC identified eight MTAs at the stringent Bonferroni significance threshold on chromosomes 1B, 2D, 3A, 3B, 3D, 6B, 7B, and 7D, with one MTA that was unmapped on the bread wheat physical map (**Supplementary Table 4**, **Supplementary Figure 3**).

The GWA scan for SDS at the individual-environment level identified considerable (45) MTAs conferring resistance to STB at different growth stages (**Supplementary Table 4**, **Supplementary Figures 4**, **5**). The analysis for disease data measured in 2015 at Holetta identified six MTAs for STB resistance at heading on chromosomes 1D, 2A, 3A, 3D, 5A, and 7D, four MTAs effective for STB resistance at the mid-maturity stage on chromosomes 1B, 3D, and 7B, with one MTA with an unknown position on the bread wheat physical map at Holetta, and nine MTAs for SPC (six at Holetta and three at Kulumsa) on chromosomes 1A, 1B, 1D, 2B, 3A, 3D, and 7B (**Supplementary Table 4**, **Supplementary Figure 4**). However, no MTA was observed for SDS data measured at the maturity stage in the same year. Likewise, the association analysis for SDS data measured in 2016 identified 26 MTAs: 4 for STB resistance at heading at Bekoji on chromosomes 3A, 3D, and 7A, with 1 MTA with an unknown position on the bread wheat physical map, 4 for maturity stage resistance on chromosomes 1D, 4A, 6A, and 7D at the same location, 18 MTAs for SPC, which were mapped to chromosomes 1A, 1B, 2B, 2D, 3B, 5B, 5D, 6A, 6B, 7A, and 7D plus 1 MTA with an unknown position on the bread wheat physical map (**Supplementary Table 4**, **Supplementary Figure 5**).

TABLE 7 | Summary of linkage disequilibrium analyses among marker pairs and the number of significant marker pairs per chromosome and genome.

| Chromosome | TNMP | r ² | Distance (Mbp) | Significant marker pairs (P < 0.01) |
|------------|---------|----------------|----------------|-------------------------------------|
| 1A | 12,475 | 0.13 | 58.05 | 4,374 (35.06) |
| 1B | 19,900 | 0.10 | 44.93 | 5,910(29.7) |
| 1D | 10,250 | 0.128 | 57.08 | 1,783 (17.40) |
| 2A | 21,200 | 0.15 | 48.03 | 7,342 (34.63) |
| 2B | 28,100 | 0.11 | 36.88 | 9,511 (33.85) |
| 2D | 14,550 | 0.21 | 57.45 | 4,771 (32.79) |
| 3A | 17,450 | 0.09 | 57.38 | 4,259 (24.41) |
| 3B | 22,000 | 0.12 | 49.85 | 6,795 (30.89) |
| 3D | 14,650 | 0.12 | 56.41 | 3,940 (26.90) |
| 4A | 13,050 | 0.11 | 75.78 | 3,830 (29.35) |
| 4B | 9,600 | 0.14 | 96.54 | 3,463 (36.08) |
| 4D | 2,700 | 0.03 | 218.80 | 1,71 (6.34) |
| 5A | 17,600 | 0.09 | 52.97 | 4,897 (27.83) |
| 5B | 22,750 | 0.14 | 40.22 | 8,193 (36.01) |
| 5D | 11,100 | 0.12 | 59.51 | 2,273 (20.48) |
| 6A | 14,700 | 0.08 | 57.51 | 3,928 (26.72) |
| 6B | 18,800 | 0.11 | 50.53 | 6,163 (32.78) |
| 6D | 7,950 | 0.05 | 87.75 | 911 (11.46) |
| 7A | 22,750 | 0.09 | 42.53 | 6,382 (28.05) |
| 7B | 22,450 | 0.11 | 42.47 | 7,492 (33.37) |
| 7D | 14,100 | 0.08 | 60.62 | 2,246 (15.93) |
| A genome | 119,225 | 0.11 | 56.04 | 35,012 (29.44) |
| B genome | 143,600 | 0.12 | 51.63 | 47,527 (33.24) |
| D genome | 75,300 | 0.11 | 65.70 | 15,184 (20.16) |
| Total | 338,125 | 0.11 | 57.79 | 97,723 (27.61) |

TNMP, total number of marker pairs; Mbp, mega-base pairs.

Putative QTL for STB resistance were identified by combining the MTAs based on their genomic positions using a window of physical distance (in Mbp) determined through a pair-wise LD analysis of the genome-wide scanned SNPs. **Supplementary Figure 6** presents a scatter plot of the genome-wise pairwise LD r^2 values between the SNPs on each chromosome against inter-marker physical distance. The MTAs falling on the same linkage group within the physical distance for LD decay specific for that chromosome were assigned to the same putative QTL. Hence, based on the LD criteria, the 53 markers were assigned to 33 putative QTLs (**Table 8, Figure 5**).

The association analysis for STB resistance at heading in the individual environments identified nine putative QTL localized on chromosomes 1D (qSTB.07), 2A (qSTB.08), 3A (qSTB.16 and qSTB.17), 3D (qSTB.20 and qSTB.21), 5A (qSTB.23), 7A (qSTB.30), and 7D (qSTB.33) (**Table 8, Figure 5**). The combined measure of the SDS at the mid-maturity stage did not reveal any QTL. However, measuring the same trait in 2015 at Holetta identified three putative QTLs on chromosomes 1B (qSTB.04), 3D (qSTB.21), and 7B (qSTB.31) (**Table 8**) that were effective for STB resistance at the mid-maturity stage. Likewise, the SDS measured at maturity in 2016 at Bekoji provided four

putative QTLs on chromosomes 1D (qSTB.05), 4A (qSTB.22), 6A (qSTB.27), and 7D (qSTB.32) (**Table 8, Figure 5**). However, the same phenotype measured across all environments did not provide any putative QTLs effective for STB resistance.

We identified seven QTLs for the SPC in the analysis of means over all the six environments: qSTB11 on 2D; qSTB15 on 3A; qSTB19 on 3B; qSTB21 on 3D; qSTB28 on 6B; qSTB31 on 7B; qSTB32 on 7D. These seven QTL modeled 9.7 to 13% of the phenotypic variation. Three of these QTLs (qSTB.11 on 2D, qSTB.14 on 3A, and qSTB.19 on 3D) were not significant in the analyses of the six environments; two (qSTB.28 and qSTB.32) were significant in one of the environments; the other two (qSTB.21 and qSTB.31) were significant in two environments (**Table 8, Figure 5**). Measuring the same trait in 2015 identified seven putative QTLs on chromosomes 1A (qSTB.01), 1B (qSTB.04), 1D (qSTB.06), 2B (qSTB.10), 3A (qSTB.17), 3D (qSTB.21), and 7B (qSTB.31) (**Table 8, Figure 5**). Similarly, the association analysis for the SPC data measured during 2016 identified effective putative QTLs on chromosomes 1A (qSTB.01 and qSTB.03), 1B (qSTB.04), 2B (qSTB.09), 2D (qSTB.12-14), 3B (qSTB.18), 5B (qSTB.24 and qSTB.25), 5D (qSTB.26), 6A (qSTB.27), 6B (qSTB.28), 7A (qSTB.29), and 7D (qSTB.32 and qSTB.33) (**Table 8, Figure 5**).

The functional association of the identified QTLs for STB resistance was further investigated by annotating genes found in the QTL regions on the recently released IWGSC RefSeq Annotation v2. The analysis resulted in several disease resistance-associated genes involved in plant defense systems (**Supplementary Table 7**). For instance, the high-confidence candidate genes *TraesCS1A02G279300* on 1A, *TraesCS1B02G332400* on 1B, *TraesCS1D02G001700* on 1D, and *TraesCS2A02G297500* on 2A are highly involved in systemic acquired resistance (SAR), which refers to the long-lasting, broad-spectrum resistance of plants to pathogen infections. Specifically, the high-confidence gene detected near *qSTB.08* on chromosome 2A (*TraesCS2A02G297500*) controls mitogen-activated protein kinase (MAPK) cascades, which are involved in signaling multiple defense responses of plants against pathogen attacks (Meng and Zhang, 2013).

MTAs for Agronomic Traits

The combined measures of the agronomic traits, such as days to heading, days to flowering, days to maturity, grain-filling duration, grain yield, and 1,000-kernel weight, did not provide any MTAs at the stringent Bonferroni significance threshold used except for plant height, which resulted in one MTA on chromosome 7A at the 514.43 Mbp position (**Supplementary Table 6**). However, the dissections of these traits in the individual environments in separate years identified considerable MTAs at the significance threshold utilized (**Supplementary Figure 7**). In particular, a GWA scan for days to heading in 2015 at Holetta provided six MTAs on chromosomes 1A, 5A, 5B, 6A, 6B, and 7B (**Supplementary Table 6**). The same trait measured in 2016 at Kulumsa identified three significant (FDR < 0.05) SNPs on chromosomes 2B, 5D, and 6A (**Supplementary Table 6**). Additionally, the total phenotypic variance explained by the SNPs for this trait ranged from

TABLE 8 | Summary of the putative QTLs identified across bread wheat chromosomes for STB resistance.

| Putative QTL | Chr | Mapposition (bp) | No.of MTAs | Flanking Markers | | | Phenotype_Location_Year | R ² |
|--------------|-----|---------------------|------------|------------------|---|------------------|--|----------------|
| | | | | Position* (Mbp) | Left/Right | Position** (Mbp) | | |
| qSTB.01 | 1A | 366278319 | 1 | 366.28 | 9766808 F 0-10:A>G-10:A>G /1087379 F 0-64:G>A-64:G>A | 374.57 | SPC_Kulumsa_2016 | 6.91 |
| qSTB.02 | 1A | 474702375 | 1 | 474.56 | 4409931 F 0-10:T>C-10:T>C/2263809 F 0-17:G>C-17:G>C | 475.71 | SPC_Kulumsa_2015 | 9.21 |
| qSTB.03 | 1A | 566369413 | 1 | 565.20 | 987669 F 0-11:T>G-11:T>G/1863565 F 0-7:G>A-7:G>A | 566.56 | SPC_Kulumsa_2016 | 5.98 |
| qSTB.04 | 1B | 558551443-587138312 | 4 | 556.11 | 3948637 F 0-11:A>G-11:A>G/1276419 F 0-52:A>C-52:A>C | 587.28 | SPC and SDS at mid-maturity | 5.98–9.67 |
| qSTB.05 | 1D | 3324483 | 1 | 2.95 | 2248863 F 0-53:G>T-53:G>T/1863120 F 0-61:T>C-61:T>C | 4.82 | SDSM_Bekoji_2016 | 273 |
| qSTB.06 | 1D | 375956648 | 1 | 363.25 | 1217216 F 0-11:G>C-11:G>C/1119123 F 0-11:C>G-11:C>G | 377.58 | SPC_Holetta_2015 | 9.66 |
| qSTB.07 | 1D | 463434850 | 1 | 462.07 | 1034027 F 0-63:C>G-63:C>G/1398976 F 0-52:G>C-52:G>C | 464.86 | SDSH_Holetta_2015 | 10.44 |
| qSTB.08 | 2A | 514858369 | 1 | 513.83 | 1181149 F 0-28:G>A-28:G>A/1102718 F 0-68:C>A-68:C>A | 21.69 | SDSH_Holetta_2015 | 10.63 |
| qSTB.09 | 2B | 243083729 | 1 | 237.98 | 100665389 F 0-10:A>G-10:A>G/3064852 F 0-14:G>A-14:G>A | 249.19 | SPC_Kulumsa_2016 | 6.63 |
| qSTB.10 | 2B | 700740191 | 1 | 698.10 | 1127049 F 0-20:T>C-20:T>C/1220715 F 0-13:A>G-13:A>G | 701.09 | SPC_Holetta_2015 | 9.84 |
| qSTB.11 | 2D | 288602370 | 1 | 215.60 | 3025921 F 0-19:G>T-19:G>T/1107980 F 0-6:T>C-6:T>C | 331.55 | SPC_Combined | 9.68 |
| qSTB.12 | 2D | 450991087 | 1 | 443.24 | 2262159 F 0-55:C>A-55:C>A/991014 F 0-5:G>C-5:G>C | 461.85 | SPC_Bekoji_2016 | 7.35 |
| qSTB.13 | 2D | 593032041 | 1 | 591.60 | 2251911 F 0-13:A>G-13:A>G/1078056 F 0-40:C>T-40:C>T | 594.54 | SPC_Bekoji_2016 | 6.01 |
| qSTB.14 | 2D | 598728762 | 1 | 595.56 | 5324627 F 0-45:G>C-45:G>C/2246647 F 0-7:T>C-7:T>C | 598.73 | SPC_Kulumsa_2016 | 7.35 |
| qSTB.15 | 3A | 8862385 | 1 | 8.74 | 2256311 F 0-9:C>G-9:C>G/1088933 F 0-37:C>T-37:C>T | 12.87 | SPC_Combined | 9.82 |
| qSTB.16 | 3A | 203418249 | 1 | 161.44 | 12470406 F 0-23:A>G-23:A>G/992022 F 0-9:G>A-9:G>A | 222.04 | SDSH_Bekoji_2016 | 2.92 |
| qSTB.17 | 3A | 710771071 | 2 | 710.34 | 989196 F 0-7:A>T-7:A>T/4989102 F 0-40:G>A-40:G>A | 711.04 | SDSH_Holetta_2015 SPC_Holetta_2015 | 9.92 |
| qSTB.18 | 3B | 17785833 | 1 | 17.11 | 1244651 F 0-19:A>G-19:A>G/998652 F 0-18:C>T-18:C>T | 18.45 | SPC_Kulumsa_2016 | 6.07 |
| qSTB.19 | 3B | 59645976 | 1 | 59.53 | 1263371 F 0-58:G>A-58:G>A/1110947 F 0-39:T>C-39:T>C | 60.37 | SPC_Combined | 9.75 |
| qSTB.20 | 3D | 42679365 | 1 | 42.63 | 981546 F 0-39:T>C-39:T>C/4911094 F 0-6:T>C-6:T>C | 45.94 | SDSH_Bekoji_2016 | 3.48 |
| qSTB.21 | 3D | 593664469 | 5 | 593.66 | 1102020 F 0-37:G>A-37:G>A/992091 F 0-53:G>C-53:G>C | 595.02 | SDSH_Holetta_2015, SDSMM_Holetta_2015, SPC_Holetta_2015, SPC_Kulumsa_2015 and SPC_Combined | 2.67– 13.01 |
| qSTB.22 | 4A | 619375783 | 1 | 619.16 | 2263956 F 0-45:T>C-45:T>C/994022 F 0-52:G>C-52:G>C | 620.75 | SDSM_Bekoji_2016 | 2.7 |
| qSTB.23 | 5A | 688359748 | 1 | 685.96 | 3938163 F 0-43:T>C-43:T>C/2278701 F 0-37:A>T-37:A>T | 689.42 | SDSH_Holetta_2015 | 10.44 |
| qSTB.24 | 5B | 487460716 | 1 | 487.44 | 1696148 F 0-16:C>G-16:C>G/2281586 F 0-67:A>G-67:A>G | 491.07 | SPC_Kulumsa_2016 | 6.92 |
| qSTB.25 | 5B | 538706298 | 1 | 538.31 | 5582250 F 0-47:T>C-47:T>C/1097026 F 0-40:C>A-40:C>A | 539.08 | SPC_Bekoji_2016 | 6.92 |
| qSTB.26 | 5D | 541603929 | 1 | 539.15 | 1696148 F 0-16:C>G-16:C>G/6038202 F 0-6:C>T-6:C>T | 541.68 | SPC_Kulumsa_2016 | 6.13 |
| qSTB.27 | 6A | 607427728-609480220 | 2 | 607.43 | 2328288 F 0-13:G>C-13:G>C/1231806 F 0-13:G>C-13:G>C | 608.28 | SPC_Kulumsa_2016 and SPC_Kulumsa_2016 | 2.72–6.01 |
| qSTB.28 | 6B | 708272196 | 2 | 706.98 | 995556 F 0-65:C>T-65:C>T/1091698 F 0-30:C>G-30:C>G | 708.02 | SPC_Combined and SPC_Bekoji_2016 | 6.13–9.91 |
| qSTB.29 | 7A | 116530515 | 1 | 116.12 | 3064815 F 0-27:A>G-27:A>G/1151957 F 0-24:T>G-24:T>G | 123.28 | SPC_Kulumsa_2016 | 7.26 |

(Continued)

TABLE 8 | Continued

| Putative QTL | Chr | Map position (bp) | No. of MTAs | Flanking Markers | | Position* (Mbp) | Position** (Mbp) | Phenotype_Location_Year | R ² |
|--------------|-----|---------------------|-------------|---|--|-----------------|------------------|---|----------------|
| | | | | Left/Right | | | | | |
| qSTB.30 | 7A | 690377106-691722567 | 2 | 3064770 F 0-20:G>A-20:G>A/3532952 F 0-38:G>A-38:G>A | | 688.71 | 690.86 | SPC_Bekoji_2016 and SDSH_Bekoji_2016 | 3.02–7.26 |
| qSTB.31 | 7B | 686989852 | 4 | 3023327 F 0-28:G>A-28:G>A/7340828 F 0-11:C>G-11:C>G | | 686.00 | 687.06 | SDSMM_Holetta_2015, SPC_Holetta_2015, SPC_Kulumsa_2015 and SPC_Combined | 8.71–9.71 |
| qSTB.32 | 7D | 21667638-63471279 | 3 | 1072451 F 0-9:C>T-9:C>T/981671 F 0-20:G>C-20:G>C | | 18.92 | 64.24 | SDSMM_Bekoji_2016, SPC_Combined and SPC_Bekoji_2016 | 2.67–10.79 |
| qSTB.33 | 7D | 528900507-531439751 | 2 | 1004225 F 0-30:G>A-30:G>A/5579572 F 0-19:G>A-19:G>A | | 526.49 | 540.62 | SPC_Bekoji_2016 and SDSH_Holetta_2015 | 6.63–11.09 |

QTL, quantitative trait locus; Chr, chromosome; MTAs, marker-trait associations; Position*, position (Mbp) of the left flanking marker; Position**, position (Mbp) of the right flanking marker; Phenotype_Location_year, Septoria disease severity measured at individual or across location levels (Holetta, Bekoji, and Kulumsa) in 2015 and 2016; R², percentage of the total phenotypic variance explained by the identified putative QTL; SDSH, Septoria disease severity at heading; SDSMM, Septoria disease severity at mid-maturity; SDSH, Septoria disease severity at maturity; SPC, Septoria progress coefficient.

8.5% for the allele 987806|F|0-16:A>G-16:A>G on chromosome 5D at 77.02 Mbp to 24.1% for the SNP1204551|F|0-57:C>T-57:C>T positioned on chromosome 1A at 499.84 Mbp (Supplementary Table 6).

The GWA scan for days to flowering in the individual environments identified six MTAs for the data measured in 2015 at Holetta on chromosomes 1A, 5A, 5B, 6A, 6B, and 7B (Supplementary Table 6, Supplementary Figure 8). The detected markers could explain 22.8–25% of the total phenotypic variation for days to flowering, while the SNP markers 1000134|F|0-15:T>C-15:T>C on chromosome 6B and 1204551|F|0-57:C>T-57:C>T on chromosome 1A accounted for the lowest and highest phenotypic variations in days to flowering in the association panel, respectively (Supplementary Table 6). Likewise, the association analysis for the days to maturity data measured in 2016 at Bekoji identified 15 MTAs pointing to nine putative QTLs on 1A, 2A, 3A, 3B, 5B, 6D, and 7A (Supplementary Table 6, Supplementary Figure 9). Three of these significant SNPs, however, were not mapped on the bread wheat physical map. The detected markers explained 11.6–15.6% of the total phenotypic variation for days to maturity, while the lowest and highest values were reported for the SNPs 2278215|F|0-18:A>G-18:A>G on 2A and for 7332831|F|0-9:T>C-9:T>C on 1A, respectively (Supplementary Table 6).

Moreover, a GWA scan on grain-filling duration data measured in 2016 revealed 15 MTAs, among which 11 were identified from the data collected at Holetta on chromosomes 1B, 2B, 3A, 3B, 6D, and 7D, plus 1 MTA with an unknown position on the bread wheat physical map (Supplementary Figure 9). Four of the significant associations for this trait were obtained from data measured at Kulumsa on chromosomes 2B, 3B, 5B, and one MTA with an unknown chromosomal position on the bread wheat genome. The total phenotypic variation for GFD explained by the SNPs ranged from 5% for the allele on chromosome 3B (5325155|F|0-26:G>T-26:G>T) to 7.9% for the SNP marker on 3A at 250.82 Mbp (5325155|F|0-26:G>T-26:G>T) (Supplementary Table 6).

The association analysis for pooled plant height data identified 1 MTA on chromosome 7A (Supplementary Figure 9) and 24 for the data measured in the individual environments (Supplementary Table 6). The plant height data measured in 2015 at Bekoji resulted in 4 MTAs on chromosomes 1A, 5A, 7A, and 7B and 10 MTAs for Kulumsa on chromosomes 1B, 2A, 5B, 5D, 6B, and 7D (Supplementary Figures 7, 8). The same trait measured in 2016 at Kulumsa provided 10 MTAs on chromosomes 3B, 5A, 5B, 6B, 6D, and 7D, with 2 MTAs that were unmapped on the bread wheat physical map (Supplementary Table 6, Supplementary Figure 9). The identified SNPs accounted for 3.3–12.3% of the total variation in plant height, and the SNP marker on 6B (SNP 2276919|F|0-10:G>T-10:G>T) at 521.99 Mbp had the largest effect (Supplementary Table 6).

The study revealed that grain yield and yield-related attributes measured over all the environments did not provide any significant SNPs. However, their association analysis based on mean values in the individual environments identified numerous MTAs. The grain yield measured in 2015 at

Kulumsa identified 10 significant MTAs on chromosomes 1A, 3A, 3D, 4A, 5A, 5D, and 7D, with one SNP that was not mapped in the bread wheat genome (**Supplementary Table 6, Supplementary Figure 8**). Similarly, a GWA scan for TKW measured in 2016 at Kulumsa provided one MTA pointing to a putative QTL on 7D at 79.52 Mbp (**Supplementary Table 6, Supplementary Figure 9**). Furthermore, the total phenotypic variations explained by the significant SNPs for yield and yield-related traits varied considerably based on traits and markers, with the lowest being 0.5% for grain yield on 3D (1045011|F|0-60:T>A-60:T>A) and the highest being 15.7% for TKW on 7D (4262368|F|0-28:A>G-28:A>G) (**Supplementary Table 6**).

The analysis revealed that some putative QTLs identified by SDS data overlapped with those determined for agronomic data. For instance, the putative QTL determined for plant height measured in 2016 at Kulumsa on chromosome 6B was co-mapped with qSTB.28, which was identified for combined SPC. Similarly, the putative QTL identified on 7B for plant height measured at Bekoji in 2015 was co-mapped with qSTB.31, which was also identified for combined SPC. In addition, the putative QTLs mapped on chromosome 7D for plant height measured in 2015 and 2016 at Kulumsa and for GFD and TKW data measured in 2016 at Holetta and Kulumsa, respectively, were co-mapped with qSTB.32, which was identified for the SDS data measured at the maturity stage at Bekoji and for pooled SPC data.

DISCUSSION

The phenotypic evaluation revealed significant genetic variability for STB resistance among the tested wheat genotypes, thus confirming the availability of relevant alleles for future breeding and improvement. The observed high broad-sense heritabilities within the individual environments ($H^2 = 0.58\text{--}0.99$) and across the environments ($H^2 = 0.72\text{--}0.81$) indicated a strong genetic signal in the data, which can be used for improving STB resistance through selection. Similarly, a high broad-sense heritability ($H^2 = 0.78$) for STB resistance was reported for European bread wheat varieties in Germany (Muqaddasi et al., 2019) and Tunisia ($H^2 = 0.55$) (Berraies et al., 2014). The present field evaluation results confirmed that STB infestation is significantly influenced by year, location, and all interaction effects, thus confirming the importance of multiple locations and years for germplasm evaluations at disease hotspots to identify durable and stable STB-resistant genotypes.

The correlation analysis revealed that Septoria disease ratings had negligible negative correlations with plant height, indicating that tallness only had a weak contribution for reducing STB infections (Muqaddasi et al., 2019). The lack of or slight negative correlations of SDS traits with the agronomic traits HD, FD, GFD, NKPS, and NKS and the moderately negative correlation with MD indicate that genotypes with late phenology could escape STB with slightly reduced infection. Moreover, the significant negative correlations of STB infection with yield and yield-related traits such as HLW, TKW, and KN could most likely be due to the fact that the infection of the flag and second leaves at the grain-filling stage could significantly influence

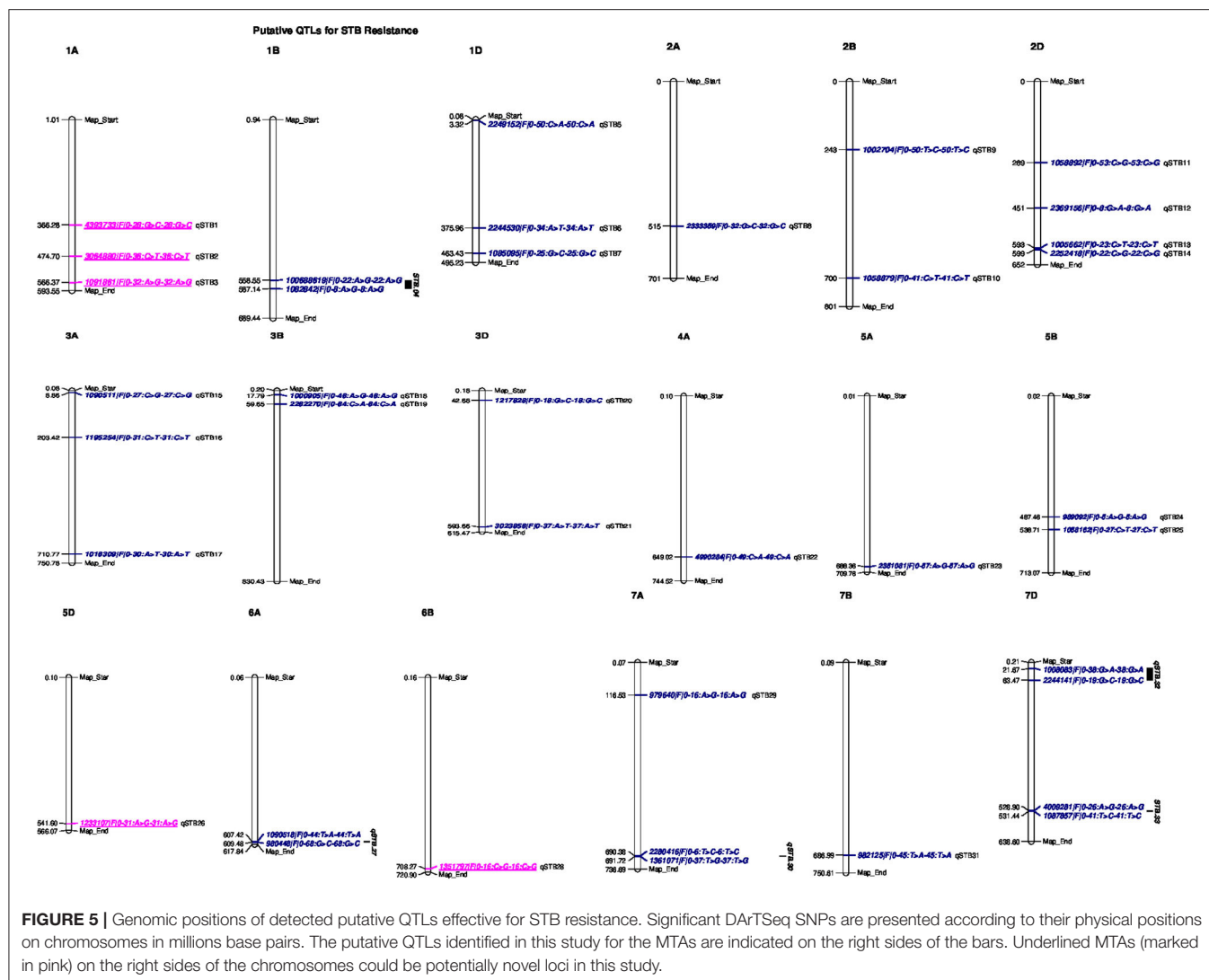
the photosynthesis process, and, thus, result in reduced grain yield. Negative associations of SDS with days to flowering, days to maturity, number of seeds per spike, and thousand-grain weights were also reported for Ethiopian durum wheat (Kidane et al., 2017).

The STRUCTURE and principal component analyses revealed population stratifications and admixtures, suggesting the need to use a powerful statistical model in the association analysis that controls for spurious marker-trait associations. The analyses suggested two sub-groups in the population. The very powerful statistical model used in the analysis, FarmCPU, sufficiently accounted for population stratification, familial relatedness, and marker effects, which consequently reduced the confounding effects that could result in false-positive MTAs. This was confirmed by visualizing the Q-Q and Manhattan plots. Similar indistinct population stratifications, higher admixtures, and weak population sub-structuring were reported among 371 European wheat genotypes based on 35k and 90k SNP marker arrays (Muqaddasi et al., 2019).

Like previous reports, this study confirmed the unequal distribution of the SNP markers among wheat genomes, where most were harbored by the A (10,317) and B genomes (10,979), while fewer SNPs (9,756) were harbored by the D genome (Berkman et al., 2013; Edae et al., 2015; Rahimi et al., 2019). This variation most likely resulted from the evolutionary and domestication history of the crop (Dvorak et al., 2006; Jordan et al., 2015), where the D genome had less time to accumulate mutations.

These analyses revealed that the LD between the markers and genes contributing to STB resistance declined to $r^2 < 0$ within a physical distance of 1.26–105.61 Mbp in all the chromosomes, with an overall mean of 31.44 Mbp. This is much lower than the average physical distance (69.1 Mbp) for LD decay in Ethiopian durum wheat at the critical threshold of $r^2 = 0.2$ (Alemu et al., 2021). The marker distances at which the LD decayed across the older sub-genomes (A and B) were relatively lower than those for the D sub-genome, most likely because of the long evolutionary history of the A and B genomes as compared with the D genome. Furthermore, the LD between alleles can decay because of a number of factors such as selection, recombination, the mating system, genetic drift, mutation, and/or population relatedness (Stich and Melchinger, 2010). Hence, it is likely that the shorter selection history of the D sub-genome did not allow linkage breakdown due to the recombination that occurs between SNPs located at longer physical distances.

The GWAS analysis identified 53 MTAs pointing to 33 QTL for STB resistance and 82 MTAs for agronomic traits where markers had a FDR $p \leq 0.05$ and a Bonferroni correction significance threshold of 0.15. The number of SDS MTAs identified in this study was significantly lower than that in the findings of Rahimi et al. (2019), who reported 313–394 MTAs for an Iranian bread wheat association panel. However, this number was still substantially higher than that in the report of Kidane et al. (2017), who identified 35 significant associations for an Ethiopian durum wheat panel. Only seven QTLs for SPC were identified in an analysis of the mean over environments, while



none was detected for SDS. Kidane et al. (2017) also reported no QTLs for SDS in an analysis of means.

More QTL were noted in the analysis of data from individual environments, although none was detected in more than four of the six environments and 24 of the 33 were detected in just one environment. The failure to detect QTL effects over environments could be due to the seasonal specificity of QTL effects, disease pressure, or the use of a very stringent FDR level that controls type I errors but leads to increased type II errors, e.g., declaring a QTL not significant when it actually is. The 2015 growing season at Holetta was characterized by extended and heavy rainfall that resulted in the highest STB natural infestations across all the growth stages. Additionally, in this growing season, 12 individual QTLs were detected (five for SPC and seven for SDS), 3 of which affected both SPC and SDS. In contrast, the heavy rainfall and prolonged moisture experienced at Bekoji in 2016 produced the highest SDS, while 12 QTLs were detected, of which 6 were for SPC, 4 for SDS, and 2 that affected both traits, using the data that were obtained. Therefore, climatic conditions,

such as persistent crop moisture and prolonged heavy rain, favor the successful infection and spread of the disease throughout the crop canopy (Fones and Gurr, 2015). Furthermore, no QTLs for SDS were detected for Kulumsa in either year. Although a total of 11 QTLs for SPC were detected in the same environment, none was repeated over the years. The different climatic conditions may have caused the later onset of the disease in both growing seasons. However, one QTL, qSTB.21, had the most repeatable effect and was significant for SPC overall for both SPC in two environments SDS at the heading and mid-maturity stages in 2015 at Holetta.

In this study, the GWA scan on SDS data measured at heading at Holetta in 2015 and at Bekoji in 2016 identified putative QTLs on chromosomes 1D, 2A, 3A, 3D, 5A, 7A, and 7D. Moreover, the association analysis for SDS at the mid-maturity stage at Holetta in 2015 reported three effective putative QTLs on chromosomes 1B, 3D, and 7B, which have not been reported for this trait so far. Likewise, dissecting the disease data measured at the maturity stage identified putative QTLs on chromosomes 1D, 4A, 6A, and

7D. Moreover, the GWA scan for SPC identified putative QTLs on chromosomes 1A, 1B, 1D, 2B, 2D, 3A, 3B, 3D, 5B, 5D, 6A, 6B, 7A, 7B, and 7D.

Although the different mapping methods, marker systems, and populations used can make it difficult to compare QTL positions from different studies, some of the QTLs detected in this analysis coincided with the mapping positions of previously reported major STB resistance genes. Hence, the putative QTL on 1B (qSTB.04) may represent *Stb2* (Liu et al., 2013) and/or *Stb11* (Chartrain et al., 2009), the QTLs on 1D (qSTB.05–07) may represent *Stb10* (Chartrain et al., 2005), the QTL on 2B may represent *Stb9* (Chartrain et al., 2009), the QTLs on 3A (qSTB.15–17) may represent *Stb6* (Brading et al., 2002) and/or *StbSm3* (Cuthbert, 2011), the QTLs on 3B may represent *Stb14* (Cowling, 2006), the QTLs on 3D may represent *Stb16q* (Tabib Ghaffary et al., 2012), the QTL on 4A may represent *Stb7* (McCartney et al., 2003) or *Stb12* (Chartrain et al., 2005), the QTLs on 5A may represent *Stb17* (Tabib Ghaffary et al., 2012), the QTLs on 5B may represent *Stb1* (Adhikari et al., 2004a), the QTL on 6A may represent *Stb15* (Arraiano et al., 2007), the QTLs on 7A may represent *Stb3* (Goodwin and Thompson, 2011) or *TmStb1* (Jing et al., 2008), the QTL on 7B may represent *Stb8* (Adhikari et al., 2003) or *Stb13* (Cowling, 2006), and the putative QTLs on 7D may represent *Stb4* (Adhikari et al., 2004b) or *Stb5* (Arraiano et al., 2001). Similar to the results of this study, Kollers et al. (2013) also reported significant MTAs for STB resistance on chromosomes 2A and 2D. The present result also agrees with Muqaddasi et al. (2019), who reported an adult-plant-stage STB resistance QTL on chromosome 4A.

The identification of defense-related candidate genes, such as *TraesCS1A02G279300*, *TraesCS1B02G332400*, *TraesCS1D02G278400*, *TraesCS2B02G233600*, *TraesCS2A02G297500*, *TraesCS2D02G497400*, *TraesCS2D02G506300*, and *TraesCS4A02G341300*, in the vicinity of the significant markers indicates the possible functional association of the detected QTL regions in plant defense systems against pathogen infections. For instance, the translations of the genes found in qSTB.02 (*TraesCS1A02G279300*) and qSTB.04 (*TraesCS1B02G332400*) on chromosomes 1A and 1B, respectively, are involved in the jasmonic acid and ethylene-dependent systemic-acquired resistance of plants to pathogen infections. This systemic acquired resistance (SAR) is a broad-spectrum, long-lasting resistance acquired after the initial localized infection of plants by pathogens (Lawton et al., 1995). Furthermore, the gene *TraesCS2A02G297500* found in qSTB.08 on chromosome 2A controls MAPK cascades, which are involved in signaling multiple defense responses, including the biosynthesis/signaling of plant stress/defense hormones, reactive oxygen species (ROS) generation, stomatal closure, defense gene activation, phytoalexin biosynthesis, cell wall strengthening, and hypersensitive response (HR) cell death. Moreover, most of the genes in proximity to the detected significant markers are inferred to be involved in salicylic acid (SA) biosynthesis, with SA being an important plant hormone that is best known for mediating host responses upon pathogen infection (Lefevre et al., 2020).

In this study, we discovered some STB resistance QTL that appear to be novel. These include the putative QTLs on chromosomes 1A (qSTB.01–3), 5D (qSTB.26), and 6B (qSTB.28) that explained >5% of the genetic variations, suggesting their relevance for wheat resistance breeding against STB. To the best knowledge of the authors, none of the known major STB resistance genes published in existing literature have been mapped to these regions of the wheat chromosomes; therefore, these QTLs could be considered novel.

Moreover, the study revealed that some of the putative STB resistance QTLs were co-located with QTL for agronomic traits. For instance, the putative QTLs derived from plant height measured in 2016 at Kulumsa on chromosome 6B ($R^2 = 11.36$) and in 2015 at Bekoji on chromosome 7B ($R^2 = 8.88$) were co-mapped with qSTB.28 and qSTB.31, which were identified for combined SPC. Likewise, the putative QTLs mapped on chromosome 7D for grain-filling duration ($R^2 = 4.69$) and 1,000-kernel weight measured in 2016 ($R^2 = 0.54$) at Holetta and Kulumsa, respectively, were co-mapped with qSTB.32, which was identified for the SDS data measured at the maturity stage at Bekoji and for the pooled SPC data. Furthermore, STB is the most destructive foliar disease in Ethiopia. Hence, the infection of the flag and second leaves, which contributes most to photosynthetic assimilates at the grain-filling stage (King et al., 1983; Muqaddasi et al., 2019), can result in the substantial loss of grain weight and yield. This is consistent with the findings of Kidane et al. (2017), who reported the co-mapping of putative QTLs for 1,000-kernel weight with SDS data. Moreover, the putative QTL on chromosome 1A identified for grain yield measured at Kulumsa in 2015 ($R^2 = 0.79$) was co-mapped with qSTB.03, which was identified for the SPC measured in the same environment. It is, therefore, expected that the vertical progression rate of the disease could affect grain yield by influencing grain filling and the number of seeds produced per spike.

In this study, most of the QTLs identified for agronomic and phenological traits did not overlap with those detected for SDS traits, likely because of the lack of common genetic effects for STB resistance and these traits. Many of the correlations of STB traits with agronomic traits were non-significant and had negligible to weak negative coefficients, indicating that the traits were independent. However, some level of co-localization was observed for the putative QTL for days to heading and days to flowering measured at Holetta in 2015 on chromosome 6B ($R^2 = 22.82$), with the putative QTL qSTB.28 on 6B being identified for the SPC measured at Bekoji in 2016 and for the pooled data.

CONCLUSIONS

In this study, the genetic architecture of adult-plant resistance to STB was explored in bread wheat using high-density, genome-wide SNP markers and multi environment-derived phenotype data. The analysis revealed that the association panel possessed considerable STB resistance alleles that could be deployed to improve wheat resistance to the prevailing *Z. tritici* populations in Ethiopia. Several genotypes with better resistance than the moderately resistant check King-bird were identified.

Furthermore, the GWAS identified 33 putative QTLs, which were associated with 53 SNPs. Most (24) of the QTLs were detected in just one environment, suggesting the presence of resistance gene/genes effective against location-specific *Z. tritici* races. The detected QTLs also explained 2.7–13.2% of the total phenotypic variance for STB resistance. Several disease resistance-associated gene/s were also identified in the proximity of the detected SNPs, which can be targeted in efforts to understand the actual causative genes at the associated loci. Additionally, most of the detected putative QTLs shared similar chromosomal positions with previously reported genes and QTLs. Among these detected alleles, five were potentially novel, accounting for >5% of STB resistance. However, the effects of these QTLs need to be validated before being deployed in MAS. Finally, we conclude that the identified stably resistant wheat genotypes and the identified QTLs can be deployed in wheat breeding programs for the development of durable and broad-spectrum-resistant varieties against *Z. tritici*.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: Dryad data repository doi: 10.5061/dryad.7sqv9s4s4.

AUTHOR'S NOTE

We confirm that the manuscript has been read and approved by all the named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

AUTHOR CONTRIBUTIONS

TM was involved in the conceptualization, methodology, statistical analysis, and writing of the original draft. CS was

involved in genotyping and editing. TH was involved in data curation and visualization. BA contributed material, performed the field experiments, and edited. CZ was involved in investigation and validation. DL was involved in reviewing. SG was involved in editing. KT was involved in methodology, funding acquisition, project administration, and supervision. All the authors approved the final version of the manuscript.

FUNDING

This study was supported financially by the Ministry of Innovation and Technology (formerly Ministry of Science and Technology) of the Federal Democratic Republic of Ethiopia.

ACKNOWLEDGMENTS

The authors are very grateful to the Kulumsa Agricultural Research Center (KARC) and Holetta Agricultural Research Center for their provision of the experimental site, germplasm, and material and technical support during the field phenotypic evaluation. We are also very grateful for the Integrated Genotyping Service and Support (IGSS) team at the Biosciences Eastern and Central Africa (BeCA-ILRI) Hub in Nairobi, Kenya, for their provision of subsidized genotyping services and relevant training. The authors would also like to extend their deep gratitude to the Institute of Biotechnology, Addis Ababa University, for the technical support during the course of the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.671323/full#supplementary-material>

REFERENCES

- Adhikari, T. B., Anderson, J. M., and Goodwin, S. B. (2003). Identification and molecular mapping of a gene in wheat conferring resistance to *Mycosphaerella graminicola*. *Phytopathology* 93, 1158–1164. doi: 10.1094/PHYTO.2003.93.9.1158
- Adhikari, T. B., Cavaletto, J. R., Dubcovsky, J., Gieco, J. O., Schlatter, A. R., and Goodwin, S. B. (2004b). Molecular mapping of the *Stb4* gene for resistance to *Septoria tritici* blotch in wheat. *Phytopathology* 94, 1198–1206. doi: 10.1094/PHYTO.2004.94.11.1198
- Adhikari, T. B., Yang, X., Cavaletto, J. R., Hu, X., Buechley, G., Ohm, H. W., et al. (2004a). Molecular mapping of *Stb1*, a potentially durable gene for resistance to *Septoria tritici* blotch in wheat. *Theor. Appl. Genet.* 109, 944–953. doi: 10.1007/s00122-004-1709-6
- Akbari, M., Wenzl, P., Caig, V., Carling, J., Xia, L., Yang, S., et al. (2006). Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theoret. Appl. Genet.* 113, 1409–1420. doi: 10.1007/s00122-006-0365-4
- Alemu, S. K., Huluka, A. B., Tesfaye, K., Haileselassie, T., and Uauy, C. (2021). Genome-wide association mapping identifies yellow rust resistance loci in Ethiopian durum wheat germplasm. *PLoS ONE* 16:e0243675. doi: 10.1371/journal.pone.0243675
- Arraiano, L. N., Balaam, P., Fenwick, C., Chapman, D., Feuerhelm, P., et al. (2009). Contributions of disease resistance and escape to the control of *Septoria tritici* blotch of wheat. *Plant Pathol.* 58, 910–922. doi: 10.1111/j.1365-3059.2009.02118.x
- Arraiano, L. S., and Brown, J. K. M. (2006). Identification of isolate specific and partial resistance to *Septoria tritici* blotch in 238 European wheat cultivars and breeding lines. *Plant Pathol.* 55, 726–738. doi: 10.1111/j.1365-3059.2006.01444.x
- Arraiano, L. S., Chartrain, L., Bossolini, E., Slatter, H. N., Keller, B., and Brown, J. K. M. (2007). A gene in European wheat cultivars for resistance to an African isolate of *Mycosphaerella graminicola*. *Plant. Pathol.* 56, 73–78. doi: 10.1111/j.1365-3059.2006.01499.x
- Arraiano, L. S., Worland, A. J., Ellerbrook, C., and Brown, J. K. M. (2001). Chromosomal location of a gene for resistance to *Septoria tritici* blotch (*Mycosphaerella graminicola*) in the hexaploid wheat 'Synthetic 6x'. *Theor. Appl. Genet.* 103, 758–764. doi: 10.1007/s001220100668

- Arzani, A., and Ashraf, M. (2017). Cultivated ancient wheats (*Triticum* spp.): a potential source of health-beneficial food products. *Comprehens. Rev. Food Sci. Food Saf.* 16, 477–788. doi: 10.1111/1541-4337.12262
- Ayele, B., Eshetu, B., Betelehem, B., Bekele, H., Melaku, D., Asnakech, T., et al. (2008). “Review of two decades of research on diseases of small cereal crops” in *Increasing Crop Production through Improved Plant Protection, Proceedings of 14th Annual Conference of Plant Protection Society of Ethiopia*, ed T. Abraham (Addis Ababa), 375–416.
- Berkman, P. J., Visendi, P., Le, H. C., Stiller, J., Manoli, S., et al. (2013). Dispersion and domestication shaped the genome of bread wheat. *Plant Biotechnol. J.* 11, 564–571. doi: 10.1111/pbi.12044
- Berraies, S., Ammar, K., Gharbi, M. S., Yahyaoui, A., and Rezgui, S. (2014). Quantitative inheritance of resistance to Septoria tritici blotch in durum wheat in Tunisia. *Chilean J. Agric. Res.* 74, 35–40. doi: 10.4067/S0718-58392014000100006
- Bradbury, P. J., et al. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Brading, P. A., Verstappen, E. C. P., Kema, G. H. J., and Brown, J. K. M. (2002). A gene-for-gene relationship between wheat and *Mycosphaerella graminicola*, the Septoria tritici blotch pathogen. *Phytopathology* 92, 439–445. doi: 10.1094/PHYTO.2002.92.4.439
- Brown, J. K. M., Chartrain, L., Lasserre-Zuber, P., and Saintenac, C. (2015). Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. *Fungal Genet. Biol.* 79, 33–41. doi: 10.1016/j.fgb.2015.04.017
- Chartrain, L., Berry, S. T., and Brown, J. K. M. (2005). Resistance of wheat line Kavkaz-K4500 L.6.A.4 to Septoria tritici blotch controlled by isolate-specific resistance genes. *Phytopathology* 95, 664–671. doi: 10.1094/PHYTO-95-0664
- Chartrain, L., Sourdille, P., Bernard, M., and Brown, J. K. M. (2009). Identification and location of *Stb9*, a gene for resistance to Septoria tritici blotch in wheat cultivars Courtot and Tonic. *Plant. Pathol.* 58, 547–555. doi: 10.1111/j.1365-3059.2008.02013.x
- Cheng, B., Gao, X., Cao, N., Ding, Y., Gao, Y., Chen, T., et al. (2020). Genome-wide association analysis of stripe rust resistance loci in wheat accessions from southwestern China. *J. Appl. Genet.* 61, 37–50. doi: 10.1007/s13353-019-00533-8
- Cowling, S. G. (2006). *Identification and mapping of host resistance genes to septoria tritici blotch of wheat* [Ph.D. thesis], University of Manitoba, Winnipeg, MB, Canada.
- Cuthbert, R. (2011). *Molecular mapping of septoria tritici blotch resistance in hexaploid wheat (Triticum aestivum L.)* [Ph.D. thesis], University of Manitoba, Winnipeg, MB, Canada.
- Dalvand, M., Zafari, D., Pari, S. M. J., Roohparvar, R., and Ghafari, S. M. T. (2018). Studying genetic diversity in *Zymoseptoria tritici*, causal agent of Septoria tritici blotch, by using ISSR and SSR markers. *J. Agr. Sci. Tech.* 20, 1307–1316. doi: 10.4067/S0718-58392018000400559
- Diversity Arrays Technology (2017). *KDCompute Plugins System*. Available online at: <https://kdcompute.seqart.net/kdcompute/plugins> (accessed June 15, 2020).
- Dvorak, J., Akhunov, E. D., Akhunov, A. R., Deal, K. R., and Luo, M. C. (2006). Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Mol. Biol. Evol.* 23, 1386–1396. doi: 10.1093/molbev/msl004
- Earl, D. A., and von Holdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Cons. Genet. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7
- Ede, E. A., Bowden, R. L., and Poland, J. (2015). Application of population sequencing (POPSEQ) for ordering and imputing genotyping-by-sequencing markers in hexaploid wheat. *G3-Genes Genomes Genet.* 12, 2547–2553. doi: 10.1534/g3.115.020362
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., 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
- 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
- Eyal, Z., and Levy, E. (1987). Variations in pathogenicity patterns of *Mycosphaerella graminicola* within *Triticum* spp in Israel. *Euphytica* 36, 237–250. doi: 10.1007/BF00730670
- Eyal, Z., and Ziv, O. (1974). The relationship between epidemics of septoria leaf blotch and yield losses in spring wheat. *Phytopathology* 64, 1385–1389. doi: 10.1094/Phyto-64-1385
- FAO (2009). *High Level Expert Forum—How to Feed the World in 2050*. Food and Agricultural Organization of the United Nations Rome.
- Fones, H., and Gurr, S. (2015). The impact of Septoria tritici blotch disease on wheat: an EU perspective. *Fungal Genet. Biol.* 79, 3–7. doi: 10.1016/j.fgb.2015.04.004
- Getinet, G., van Ginkel, M., Temesgen, K., Mintwab, H., Rebeka, D., Bainbridge, A., et al. (1990). “Wheat disease survey in Ethiopia in 1988,” in *Sixth Regional Wheat Workshop for Eastern, Central and Southern Africa CIMMYT*, eds D. G. Tanner, M. van Ginkel, and W. M. Mwangi (Mexico), 153–165.
- Ghaneie, A., Mehrabi, R., Safaie, N., Abrinbana, M., Saidi, A., and Aghaei, M. (2011). Genetic variation for resistance to septoria tritici blotch in Iranian tetraploid wheat landraces. *Eur. J. Plant Pathol.* 132, 191–202. doi: 10.1007/s10658-011-9862-7
- 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
- Goodwin, S. B., and Thompson, I. (2011). Development of isogenic lines for resistance to Septoria tritici blotch in wheat. *Czech J. Genet. Plant Breed* 47, S98–S101. doi: 10.17221/3262-CJGPB
- Huang, C., Nie, X. H., Shen, C., You, C. Y., Li, W., Zhao, W. X., et al. (2017). Population structure and genetic basis of the agronomic traits of upland cotton in China revealed by a genome-wide association study using high-density SNPs. *Plant Biotechnol. J.* 15, 1374–1386. doi: 10.1111/pbi.12722
- International Wheat Genome Sequencing Consortium (IWGSC) (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191
- Jamil, M., Ali, A., Gul, A., Ghafoor, A., Napar, A. A., Ibrahim, A. M. H., et al. (2019). Genome-wide association studies of seven agronomic traits under two sowing conditions in bread wheat. *BMC Plant Biol.* 19:149. doi: 10.1186/s12870-019-1754-6
- Jing, H. C., Lovell, D., Gutteridge, R., Jenk, D., Korniyukhin, D., Mitrofanova, O. P., et al. (2008). Phenotypic and genetic analysis of the *Triticum monococcum* – *Mycosphaerella graminicola* interaction. *New Phytol.* 179, 1121–1132. doi: 10.1111/j.1469-8137.2008.02526.x
- Jordan, K. W., Wang, S., Lun, Y., Gardiner, L. J., MacLachlan, R., Hucl, P., et al. (2015). Consortium IWGS. A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biol.* 16:48. doi: 10.1186/s13059-015-0606-4
- Kankwatsa, P., Singh, D., Thomson, P. C., Babiker, E. M., Bonman, J. M., Newcomb, M., et al. (2017). Characterization and genome-wide association mapping of resistance to leaf rust, stem rust and stripe rust in a geographically diverse collection of spring wheat landraces. *Mol. Breed.* 37:113. doi: 10.1007/s11032-017-0707-8
- Kidane, Y. G., Hailemariam, B. N., Mengistu, D. K., Fadda, C., Pè, M. E., and Dell’Acqua, M. (2017). Genome-Wide Association Study of Septoria tritici blotch resistance in Ethiopian durum wheat landraces. *Front. Plant Sci.* 8:1586. doi: 10.3389/fpls.2017.01586
- King, J. E., Jenkins, J. E. E., and Morgan, W. A. (1983). The estimation of yield losses in wheat from severity of infection by *Septoria* species. *Plant Pathol.* 32, 239–249. doi: 10.1111/j.1365-3059.1983.tb02831.x
- Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., et al. (2013). Whole genome association mapping of fusarium head blight resistance in European Winter Wheat (*Triticum aestivum* L.). *PLoS One* 8:e57500. doi: 10.1371/journal.pone.0057500
- Kopelman, N. M., et al. (2015). CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Res.* 15, 1179–1191. doi: 10.1111/1755-0998.12387
- Lawton, K., Weymann, K., Friedrich, L., Vernooij, B., Uknes, S., and Ryals, J. (1995). Systemic acquired resistance in *Arabidopsis* requires

- salicylic acid but not ethylene. *Mol. Plant-Microbe Interact.* 8, 863–870. doi: 10.1094/MPMI-8-0863
- Lefeverre, H., Bauters, L., and Gheysen, G. (2020). Salicylic acid biosynthesis in plants. *Front. Plant Sci.* 11:338. doi: 10.3389/fpls.2020.00338
- 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
- Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative Usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12:e1005767. doi: 10.1371/journal.pgen.1005767
- Liu, Y. L., Zhang, I. A., Thompson, S. B., and Goodwin, and, H. W., Ohm (2013). Molecular mapping re-locates the *Stb2* gene for resistance to Septoria tritici blotch derived from cultivar Veranopolis on wheat chromosome 1BS. *Euphytica* 190, 145–156.
- Long, L., Yao, F., Yu, C., Ye, X., Cheng, Y., Wang, Y., et al. (2019). Genome-Wide Association Study for adult-plant resistance to stripe rust in Chinese wheat landraces (*Triticum aestivum* L.) from the Yellow and Huai river valleys. *Front. Plant Sci.* 10:596. doi: 10.3389/fpls.2019.00596
- Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., Jakobsen, K. S., et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345:1250092. doi: 10.1126/science.1250092
- Mathew, I., Shimelis, H., Shayanowako, A. T., Laing, M., and Chaplot, V. (2019). Genome-wide association study of drought tolerance and biomass allocation in wheat. *PLoS ONE* 14:e0225383. doi: 10.1371/journal.pone.0225383
- McCartney, H. A., Foster, S. J., Fraaije, B. A., and Ward, E. (2003). Molecular diagnostics for fungal plant pathogens. *Pest Manag. Sci.* 59:129–142. doi: 10.1002/ps.575
- Mekonnen, T., Haileselassie, T., Kaul, T., Sharma, M., Geleta, A., and Tesfaye, K. (2019). Molecular screening of *Zymoseptoria tritici* resistance genes in wheat (*Triticum aestivum* L.) using tightly linked simple sequence repeat markers. *Eur. J. Plant Pathol.* 155, 593–614. doi: 10.1007/s10658-019-01795-y
- Mekonnen, T., Haileselassie, T., Goodwin, S. B., and Tesfaye, K. (2020). Genetic diversity and population structure of *Zymoseptoria tritici* in Ethiopia as revealed by microsatellite markers. *Fungal Genet. Biol.* 141:103413. doi: 10.1016/j.fgb.2020.103413
- Meng, X., and Zhang, S. (2013). MAPK cascades in plant disease resistance signaling. *Annu. Rev. Phytopathol.* 51, 245–266. doi: 10.1146/annurev-phyto-082712-102314
- Mengistu, H., Getaneh, W., Yeshe, A., Rbka, D., and Ayele, B. (1991). “Wheat pathology research in Ethiopia,” in *Wheat Research in Ethiopia: A Historical Perspective IAR/CIMMYT*, eds G. Hailu, D. G., Tanner, and H. Mengistu (Addis Ababa), 173–217.
- Muqaddasi, Q. H., Zhao, Y., Rodemann, B., Plieske, J., Ganai, M. W., and Roder, M. S. (2019). Genome-wide association mapping and prediction of adult stage *Septoria tritici* blotch infection in European winter wheat via high-density marker arrays. *Plant Genome* 12:180029. doi: 10.3835/plantgenome2018.05.0029
- Odilbekov, F., Armonien, R., Koc, A., Svensson, J., and Chawade, A. (2019). GWAS-assisted genomic prediction to predict resistance to Septoria tritici blotch in Nordic winter wheat at seedling stage. *Front. Genet.* 10:1224. doi: 10.3389/fgene.2019.01224
- Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multi-locus genotype data. *Genetics* 155, 945–959. Available online at: <https://www.genetics.org/content/155/2/945>
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation Statistical Computing. Available online at: www.r-project.org
- Rafalski, J. A. (2010). Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13, 174–180. doi: 10.1016/j.pbi.2009.12.004
- Rahimi, Y., Bihanta, M. R., Taleei, A., Alipour, H., and Ingvarsson, P. K. (2019). Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. *BMC Plant Biol.* 19:541. doi: 10.1186/s12870-019-2165-4
- Ramadas, S., Kumar, K. T. M., and Singh, G. P. (2019). *Wheat Production in India: Trends and Prospects* London: IntechOpen.
- Rashid, Z., Singh, P. K., Vemuri, H., Zaidi, P. H., Prasanna, B. M., and Nair, S. K. (2018). Genome-wide association study in Asia-adapted tropical maize reveals novel and explored genomic regions for sorghum downy mildew resistance. *Sci. Rep.* 8:366. doi: 10.1038/s41598-017-18690-3
- Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* 8:e66428. doi: 10.1371/journal.pone.0066428
- Robinson, H. F., Comstock, R. E., and Harvey, P. H. (1949). Estimates of heritability and the degree of dominance in corn. *Agronomy Journal*, 41, 353–359
- SAS Institute Inc. (2008). *SAS/STAT® 9.2 User's Guide*. Cary, NC: SAS Institute Inc.
- Sharma, R. C., and Duveiller, E. (2007). Advancement toward new spot blotch resistant wheat in south Asia. *Crop Sci.* 47, 961–968. doi: 10.2135/cropsci2006.03.0201
- Stich, B., and Melchinger, A. E. (2010). “An introduction to association mapping in plants,” in *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* (Wallingford: CBAI), 5, No. 039. doi: 10.1079/PAVSNNR20105039
- Tabib Ghaffary, S. M., Faris, J. D., Friesen, T. L., Visser, R. G. F., van der Lee, T. A. J., Robert, O., et al. (2012). New broad-spectrum resistance to Septoria tritici blotch derived from synthetic hexaploid wheat. *Theor. Appl. Genet.* 124, 125–142. doi: 10.1007/s00122-011-1692-7
- Takele, A., Lencho, A., Getaneh, W., Hailu, E., and Kassa, B. (2015). Status of wheat Septoria leaf blotch (*Septoria tritici* Roberge in Desmaz) in South West and Western Shewa Zones of Oromiya Regional State, Ethiopia. *Res. Plant Sci.* 3, 43–48. doi: 10.12691/PLANT-3-3-1
- Tang, Y., Liu, X., Wang, J., Li, M., Wang, Q., Tian, F., et al. (2016). GAPIT Version 2: an enhanced integrated tool for genomic association and prediction. *Plant Genome* 9:1–9. doi: 10.3835/plantgenome2015.11.0120
- Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Sour. Softw.* 3:731. doi: 10.1101/005165
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980
- Wang, S., Xu, S., Chao, S., Sun, Q., Liu, S., and Xia, G. (2019). A Genome-Wide Association study of highly heritable agronomic traits in durum wheat. *Front. Plant Sci.* 10:919. doi: 10.3389/fpls.2019.00919
- Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS ONE* 14:e0208217. doi: 10.1371/journal.pone.0208217
- Xiao, Y., Liu, H., Wu, L., Warburton, M., and Yan, J. (2017). Genome-wide association studies in maize: praise and stargaze. *Mol. Plant* 10, 359–374. doi: 10.1016/j.molp.2016.12.008
- Yao, F., Zhang, X., Ye, X., Li, J., Long, L., Yu, C., et al. (2019). Characterization of molecular diversity and genome-wide association study of stripe rust resistance at the adult plant stage in Northern Chinese wheat landraces. *BMC Genet.* 20:38. doi: 10.1186/s12863-019-0736-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.

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 Mekonnen, Sneller, Haileselassie, Ziyomo, Abeyo, Goodwin, Lule and Tesfaye. 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.



Associated SNPs, Heritabilities, Trait Correlations, and Genomic Breeding Values for Resistance in Snap Beans (*Phaseolus vulgaris* L.) to Root Rot Caused by *Fusarium solani* (Mart.) f. sp. *phaseoli* (Burkholder)

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Juan Jose Ferreira,
Servicio Regional de Investigación y
Desarrollo Agroalimentario (SERIDA),
Spain

Andrés J. Cortés,
Colombian Corporation
for Agricultural Research
(AGROSAVIA), Colombia

*Correspondence:

James R. Myers
james.myers@oregonstate.edu

[†] 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: 19 April 2021

Accepted: 23 August 2021

Published: 28 September 2021

Citation:

Huster AR, Wallace LT and
Myers JR (2021) Associated SNPs,
Heritabilities, Trait Correlations,
and Genomic Breeding Values
for Resistance in Snap Beans
(*Phaseolus vulgaris* L.) to Root Rot
Caused by *Fusarium solani* (Mart.)
f. sp. *phaseoli* (Burkholder).
Front. Plant Sci. 12:697615.
doi: 10.3389/fpls.2021.697615

Abigail R. Huster^{1†}, Lyle T. Wallace^{2†} and James R. Myers^{1*}

¹ Department of Horticulture, Oregon State University, Corvallis, OR, United States, ² USDA-ARS, Plant Germplasm
Introduction and Testing Research Unit, Washington State University, Pullman, WA, United States

Root rot is a major constraint to snap bean (*Phaseolus vulgaris*) production in the United States and around the world. Genetic resistance is needed to effectively control root rot disease because cultural control methods are ineffective, and the pathogen will be present at the end of one season of production on previously clean land. A diversity panel of 149 snap bean pure lines was evaluated for resistance to *Fusarium* root rot in Oregon. Morphological traits potentially associated with root rot resistance, such as aboveground biomass, adventitious roots, taproot diameter, basal root diameter, deepest root angle, shallowest root angle, root angle average, root angle difference, and root angle geometric mean were evaluated and correlated to disease severity. A genome wide association study (GWAS) using the Fixed and random model Circulating Probability Unification (FarmCPU) statistical method, identified five associated single nucleotide polymorphisms (SNPs) for disease severity and two SNPs for biomass. The SNPs were found on Pv03, Pv07, Pv08, Pv10, and Pv11. One candidate gene for disease reaction near a SNP on Pv03 codes for a peroxidase, and two candidates associated with biomass SNPs were a 2-alkenal reductase gene cluster on Pv10 and a Pentatricopeptide repeat domain on Pv11. Bean lines utilized in the study were ranked by genomic estimated breeding values (GEBV) for disease severity, biomass, and the root architecture traits, and the observed and predicted values had high to moderate correlations. Cross validation of genomic predictions showed slightly lower correlational accuracy. Bean lines with the highest GEBV were among the most resistant, but did not necessarily rank at the very top numerically. This study provides information on the relationship of root architecture traits to root rot disease reaction. Snap bean lines with genetic merit for genomic selection were identified and may be utilized in future breeding efforts.

Keywords: common bean, disease resistance, genome wide association studies, genomic prediction, best linear unbiased prediction, root morphology, genomic selection

INTRODUCTION

Root rot is a serious disease that affects common beans (*Phaseolus vulgaris*) wherever they are grown. It has been and continues to be a primary yield limitation in both snap and dry bean production. Root rot is a broad term that can refer to infection by a variety of pathogens or complexes thereof (Abawi et al., 1985). The most serious and widespread causal pathogen, *Fusarium solani* f. sp. *phaseoli*, has been reported to cause yield losses of up to 84% in the United States (Schneider et al., 2001). This organism is the primary, although not necessarily exclusive, root rot pathogen in Oregon snap bean fields. There is currently no satisfactory management technique to control root rot in snap beans with cultural and chemical methods having met with limited success (Burke and Miller, 1983). The best cultural option available to control root rot is crop rotation but the four-to-five-year interval that is required is impractical for most farmers. With so few options, genetic resistance is of paramount importance. The benefits of resistance extend beyond mitigating disease. Without functional root systems, it is impossible to select for other traits, such as abiotic stress and nutrient use efficiency that are needed to combat climate change and adapt to agricultural intensification.

Most prior genetic analyses of *F. solani* root rot resistance have been conducted with biparental dry bean populations (Supplementary Table 1). Many were conducted with RAPD marker systems that are difficult to rectify with contemporary SNP-based maps (Chowdhury et al., 2002; Román-Avilés and Kelly, 2005; Navarro et al., 2008; Schneider et al., 2001). SNPs have become the preferred marker for linkage and association mapping because of their abundance, repeatability and reference to physical location within the genome (Blair et al., 2013; Cortés et al., 2011). Unlike others who focused exclusively on dry beans, Navarro et al. (2008) and Hagerty et al. (2015) used snap x dry bean populations to map QTL for root rot resistance. In all cases, resistance was inherited quantitatively with one to 15 QTL explaining from five to 53% of total phenotypic variance. Where reported, heritabilities have ranged from 10 to 99%, with the majority being in the low to moderate range. One genome wide association study (GWAS) has been conducted in dry bean for resistance to *F. solani* root rot. This study identified SNP associations in Andean and Middle American diversity panels (Zitnick-Anderson et al., 2020). They found sixteen unique SNP associations in an Andean diversity panel on Pv01, Pv02, Pv03, Pv04, Pv07, Pv08, Pv09, and Pv11, and seven unique SNP associations in a Middle American panel on Pv01, Pv03, Pv04, Pv07, and Pv08 (Zitnick-Anderson et al., 2020). Further GWAS studies have been conducted on root rot caused by *Pythium* spp., *Pythium ultimum*, *Fusarium oxysporum* and *Rhizoctonia solani* in dry bean (Oladzad et al., 2019; Dramadri et al., 2020; Diaz L. M. et al., 2021; Paulino et al., 2021). With *F. solani*, the studies listed in Supplementary Table 1 have not found major QTL associated with resistance and the general consensus is that resistance is conditioned by several to many genes with small individual effect.

There is evidence that the genetic background of snap beans has unique characteristics which warrants examination on its own (Wallace et al., 2018). In particular, the genetic background of snap beans is highly mixed between the Andean and Middle American gene pools with unknown

effects on the interactions of genes. Moreover, snap beans have been selected for succulent, low fiber pods mostly in isolation from dry beans since their assimilation by Europeans starting in the 1500's and this time frame may have been sufficient for unique resistance traits to evolve within snap beans.

The traditional GWAS model is a mixed linear model with a correction for kinship and population structure that adequately controls type I statistical errors. Last-generation GWAS models, such as FarmCPU, have improved sensitivity and statistical power with similar control of type I statistical errors and much improved control of type II statistical errors (Liu et al., 2016; López-Hernández and Cortés, 2019). Work on last-generation GWAS models (FarmCPU, BLINK, and SUPER) indicates that they are comparable and complement with each other when used in parallel, although subtle differences have been found, such as non-redundant results (FarmCPU) or a greater number of associated SNPs (BLINK) in a study of heat stress in common bean (López-Hernández and Cortés, 2019). Both BLINK and FarmCPU iteratively utilize a random and fixed model and may have an advantage over SUPER in having a lower type II statistical error rate (López-Hernández and Cortés, 2019).

Marker assisted selection (MAS) have been most successfully applied to traits conditioned by major genes, or in some cases, major QTL (Assefa et al., 2019) and specifically in breeding programs to introgress disease resistance. Over 40 SCAR or SRAP markers linked to resistance to 11 pathogens are available in common bean (BIC, 2021b). Only two of these are for root pathogens (*Fusarium oxysporum* and *Pythium ultimum*), where resistance is conditioned by major genes. Some studies on *F. solani* resistance indicate that the markers that were discovered may be useful in breeding for resistance. However, there is little evidence of their application in breeding programs. The underlying reason for this is probably the polygenic nature of *F. solani* resistance. MAS has not proven to be very effective for such traits. Genomic selection (GS) is emerging in common bean as a technique that allows selection of quantitative traits without the labor-intensive approach that traditional MAS would require (Assefa et al., 2019). GS models generally use many markers distributed across the genome, and as a result, are more effective than traditional MAS in selection for traits with many genes with small effect. GS has been applied to common bean for root rot (Diaz L. M. et al., 2021) as well as to agronomic traits (Keller et al., 2020), cooking time (Diaz S. et al., 2021), and nematode resistance (Wen et al., 2019; Shi et al., 2021) to discover genotypes with the best breeding values for recombination schemes, but deployment in breeding programs is only beginning.

Differing models for genomic selection are similar in their predictive accuracy. One study of maize traits found that rrBLUP had a slightly higher predictive accuracy in comparison to four other genomic prediction models (Riedelsheimer et al., 2012). Other research into genomic selection models in barley and wheat found no differences, but a study of loblolly pine found rrBLUP lacking when applied to oligogenic traits with a few major genes (Heslot et al., 2012; Resende et al., 2012).

The purpose of this research was to improve the understanding of the genetics underlying resistance to *F. solani* sp. *phaseoli* in snap beans under field conditions typically found

in a major snap bean growing region of the United States. As the genetic background of snap beans is unique, this is an important gap that needs investigation separate from previous dry bean studies of *Fusarium* root rot genetic architecture. To achieve this goal, three research focus areas were identified: (1) Analysis of root and plant morphological traits in a diversity panel of snap beans as related to root rot resistance or susceptibility, (2) GWAS on root rot resistance in a diversity panel of snap beans, and (3) Genomic prediction of cultivars to identify lines with superior breeding potential based on the totality of all marker effects in order to better capture minor allelic effects that may be missed by GWAS.

MATERIALS AND METHODS

Study Site and Experimental Design

In this study, 149 pure lines of the Common Bean Coordinated Agricultural Project (BeanCAP) Snap Bean Diversity Panel (SBDP; see data availability statement for details on this panel) were evaluated for resistance to root rot, which primarily consists of *F. solani* in Oregon. This diversity panel contains pure line examples of both centers of domestication with a representative cross section of historical and contemporary snap beans, but no wild materials. About 83% of the lines in the SBDP are of Andean center of domestication with the remainder being of Mesoamerican derivation (Wallace et al., 2018). They can be further classified into eight groups based on Structure analysis, with some lines having genetic contributions from as many as seven groups. Since snap beans have undergone a high level of intermixing relative to dry beans between the centers of domestication (Wallace et al., 2018), more than 50% of the snap beans in the panel contain some genetic background from both centers of domestication.

Strongly root rot susceptible ('Seabiscuit', 'Shade', and 'Zodiac') and strongly resistant ('Black Valentine', 'Impact', and 'Widusa') cultivars were included. The OSU cultivars included in the panel were bred and selected on the research farm under constant root rot pressure, and as a result, have high levels of resistance, and consistently grouped with the most resistant lines in the diversity panel. Additionally, the panel included 'FR-266', an experimental snap bean line bred in the Pacific Northwest for *F. solani* root rot resistance (Silbernagel, 1987). This line has been used in biparental mapping population studies of root rot resistance (Schneider et al., 2001). It has been a check in our root rot breeding nursery trials, where it shows moderate levels of resistance. The complete panel was used, except for 'BBL 274', which was unavailable for planting. In late spring of 2014 and 2015, four replicates of the SBDP were planted at the Oregon State University Vegetable Research Farm. The Vegetable Research Farm is located in Corvallis, Oregon on Chehalis silty clay loam soil at latitude N44.571209, longitude W123.243261 at 77 masl. The studies took place in our root rot "purgatory plot" that had been planted continually with snap beans for over 25 years in an effort to build a heavy pathogen population and increase disease pressure for more effective screening. In monitoring of bean root pathogens present at the Vegetable Research Farm,

we have always found *F. solani* to be the primary pathogen (see Cirak and Myers, 2021 for latest assay). To further encourage heavy and uniform disease pressure, the trials were well irrigated (2.5 cm of water weekly by solid set sprinklers) in the beginning of each season, as high soil moisture levels aid in infection. After pod set, irrigation was reduced to increase abiotic stress levels. The late season irrigation schedules were determined based on weather conditions.

The trials were planted with a modified randomized complete block design with the field divided into four replicated blocks on a north-south axis. This method of blocking was chosen as the size of this experiment exceeded previous years' plantings and extended into soil that may have had a lower level of disease pressure. Due to their unique characteristics and need for a trellis system, the pole beans were planted in a separate four block randomization at the west end of the field. The plots were 3.0 m long, planted in a single row at a density of 50 seeds per plot. Rows were spaced 75 cm apart. A border row of OSU5446, a root rot susceptible experimental line, was planted on the north and south edges of the field, as well as 1.5 m end plots on the east and west ends of each row to minimize edge effects. Planting dates were 10 June in 2014 and 21 May in 2015. The seed was treated with captan pre-emergent fungicide (Bonide Products Inc.) prior to planting to improve germination and emergence uniformity and reduce differences in stand among lines.

Field Evaluation

Data collection began when the earliest lines were at 50% buckskin pod stage (when half the pods per bush have lost their chlorophyll and have taken on a flexible, leathery texture). Each plot was evaluated at this uniform phenological stage. A Shovelomics protocol (Lynch and Brown, 2001, 2013) was used to perform evaluations. The SBDP was evaluated for several morphological traits including taproot diameter, largest basal root diameter, deepest and shallowest basal root angles, and aboveground biomass to investigate correlations between plant structure and disease resistance. Five consecutive plants from the center of the plot were dug with a 30 cm radius of soil around the roots, and carefully shaken and washed to remove the soil without damaging the roots. The five plants were evaluated on a 1–5 (1 = least and 5 = most biomass) scale as a single unit for aboveground biomass (**Supplementary Figure 1**). A subsample of two randomly selected plants from the original five were evaluated independently for taproot diameter, largest basal root diameter, deepest basal root angle, shallowest basal root angle, adventitious roots (1–3 scale; 1 = few, 3 = many roots), and disease severity (1–5 scale; **Table 1**). In evaluating disease severity of *F. solani*, nearly all researchers have used 1–5 or 1–9 visual rating scales (Azzam, 1956; Baggett et al., 1965; Abawi, 1990; Hagerty et al., 2015; BIC, 2021a). Taproot and largest basal root diameter were recorded with digital calipers. The measurements were taken 1 cm. below where the root emerged from the hypocotyl. The deepest (closest to the taproot) and shallowest (closest to the soil line) basal root angles were measured by laying the specimen on a cutting board marked with protractor angle increments (**Supplementary Figure 2**).

TABLE 1 | Scale for rating *Fusarium solani* root rot symptoms in the BeanCAP Snap Bean Diversity Panel grown at the Oregon State University Vegetable Research Farm for a genome wide association study.

| Score | Root rot rating scale description |
|-------|--|
| 1.0 | Clean white root |
| 1.5 | Few external red or brown lesions |
| 2.0 | Some external lesions, but root still firm and white inside |
| 2.5 | Some external lesions, red discoloration of pith, but root is firm |
| 3.0 | Significant external infection, red to brown pith |
| 3.5 | Spongy brown lesions are present |
| 4.0 | Root is soft and rotten |
| 4.5 | Root is very rotten, falling off |
| 5.0 | Root is absent, plant ends in rotten stump |

Root angle difference, root angle average, and root angle geometric mean were calculated from deepest and shallowest root measurements. Root angle difference was the shallowest root angle subtracted from the deepest root angle. This conveys the span of the soil profile accessed by the plant. Root angle average is the mean of the deepest and shallowest root angles and expresses the general orientation of the roots, from zero to 90°. Root angle geometric mean is the geometric mean of the root angle average and the root angle difference and was formulated to provide a single value that integrated soil profile span and root orientation.

Statistical Analysis of Field Trials

To characterize the variation observed in the 2014 and 2015 trials, the following statistical approach was used. First, homogeneity of variances across years was examined using PROC GLIMMIX (SAS version 9.3: SAS institute, Cary, NC) using the model [Trait] = Variety Rep(Year) Year Variety*Year with Year treated as a random effect and the Covtest option to test for homogeneity of variances. Variances from 2014 and 2015 demonstrated homogeneity, and both years of data were combined into a single analysis. Second, normality by year was examined using PROC GLM with the model [Trait] = Rep Variety. Third, a mixed model analysis of variables with years combined was performed using PROC GLM with the model [Trait] = Variety Year Rep(Year) Year*Variety with Year, Rep(Year) and Year*Variety treated as random effects. As the two individual plants measured from each plot were intended to capture information on a plot-mean basis rather than an individual plant basis, mean scores for each plot were used.

Multiple Correlation Analysis Among Traits

To evaluate whether root morphological traits and disease severity were positively or negatively associated, a Pearson's correlation coefficient analysis was performed in SAS 9.3 on the least square means of the phenotypic data for disease severity, aboveground biomass, adventitious roots, basal root diameter, taproot diameter, shallowest root angle, deepest root angle, root angle difference, root angle average, and root angle geometric mean. Least square means were generated from combined data from the 2014 and 2015 trials when ANOVAs were conducted

as described above. Correlations were generated for all pairwise combinations of traits.

Genotyping

The genotypic dataset consisted of 10,607 SNPs generated by using two Illumina iSelect 6K Gene Chip sets (BARCBEAN6K_1 and BARCBEAN6K_2) (Song et al., 2015). These BeadChips were designed following sequencing a diverse set of 17 dry bean cultivars with 10 from the Mesoamerican and seven from the Andean centers of domestication. SNPs with 50% or greater missing data were discarded (Song et al., 2015). Remaining missing genotypes were imputed using fastPHASE, which uses the Hidden Markov Model to indicate the cluster membership of haplotypes (Scheet and Stephens, 2006). Genotypic data for the 'Panama' genotype was unavailable and was excluded from the GWAS and BLUP analysis.

Heritability

Narrow sense and broad sense heritability are essentially equivalent in a highly inbred crop such as common bean. With complete homozygosity, it can be assumed that there are no dominance effects present. In the absence of dominance effects, variance among inbred lines, or Var(G), provides an estimate of additive genetic variance or Var(A), rendering the two equations equivalent (Hallauer et al., 2010). Additive x additive epistasis may inflate estimates of narrow sense heritability, but is typically minimal in a diploid crop such as common bean. The formula:

$$\hat{h}^2 = \frac{\hat{\sigma}_g^2}{\hat{\sigma}_{re}^2 + \frac{\hat{\sigma}_{ge}^2}{e} + \hat{\sigma}_g^2}$$

was used to determine heritability, where $\hat{\sigma}_g^2$ is the estimated genotypic variance component, $\hat{\sigma}_{ge}^2$ is the estimated genotype by environment interaction variance component, $\hat{\sigma}^2$ is the estimated experimental error variance, e is the number of environments, and r is the number of replications per environment. Heritability for each trait was calculated using SAS code developed by Holland et al. (2003). Mixed model analysis (PROC MIXED, SAS 9.3) was used to obtain variance components. Variance components were estimated using the restricted maximum likelihood (REML) method. All model components were treated as random effects. Heritability was calculated on a line mean basis.

Genome-Wide Association Study

The entire SNP dataset was utilized for GWAS analysis. The phenotypic data used for GWAS was a single value for each trait, averaged across four reps and two years. Due to the incongruity of a pole bean plant architecture for biomass measurements, pole type beans were removed from the biomass analysis leaving 139 genotypes (lines) for this analysis. All other traits were measured with the full set of genotypes.

The FarmCPU statistical method was performed in version 4.0.2 of the R software environment (Liu et al., 2016). To derive SNP R^2 values, FarmCPU was run within GAPIT (version 3) with the added code, Random.model = TRUE. The SNP data was

formatted in Microsoft Excel and was filtered for a minor allele frequency (MAF) of 0.05 within R.

The principal component analysis (PCA) was conducted in TASSEL, version 5.2.73.¹ Principal components one to five accounted for 22, 33, 41, 48, and 52% of the variation, respectively. Based on the widely accepted criterion of principal components accounting for between 25 and 50% of the variation (Oladzad et al., 2019; Zitnick-Anderson et al., 2020), the choice of principal components was narrowed to between two and four. To further narrow the choice of principal components, QQ plots were examined for fit around the null distribution to make the final selection of two principal components (**Supplementary Figure 3**). Linkage Disequilibrium (LD) heat maps for individual chromosomes were also generated in TASSEL using the full matrix in lieu of the sliding window.

Two different thresholds were examined for a cutoff of significance in the Manhattan plots. The more conservative threshold was a Bonferroni cutoff that utilized the effective marker number of 2,411 as determined by the SimpleM method (Gao et al., 2010). This generated an alpha 0.05 threshold of 4.68 as expressed as a negative log value. In addition, a 10,000 bootstrap threshold was generated for an alpha of 0.05 (Mamidi et al., 2014). This bootstrap identified a threshold of 4.51 negative log.

Candidate Gene Search

Associated SNP positions were located in the common bean genome as shown in the Phytozome JBrowse genome browser (Phytozome, version 12.1; *P. vulgaris* genome, version 2.1). Using conservative estimates of linkage disequilibrium in common bean (Soltani et al., 2018; Oladzad et al., 2019) and in consideration of the fact that no wild materials are included in our panel, we chose to bracket a region of ± 100 kb in our search for candidate genes. Each gene model within the bracketed region was researched for its potential role in disease resistance or biomass.

Genomic Prediction

Genomic estimated breeding values were calculated by adding the fixed effect BLUE value for a given trait to the random effect BLUP value for a given bean line and trait as determined by the rrBLUP R package (Endelman, 2011). rrBLUP is equivalent to gBLUP when QTLs are many, there are no major QTLs and QTLs are evenly distributed across the genome (Bernardo, 2020). They differ in that rrBLUP calculates SNP effects from a set of related individuals whereas gBLUP uses markers to estimate relatedness among individuals. Genomic prediction utilized the entire SNP dataset.

To evaluate the predictive power of the rrBLUP calculations, cross validation was performed by randomly splitting all the genotypes within this study into a training set and validation set. The models evaluated used ratios of training set to validation set of 60:40, 70:30, 80:20, and 90:10%. Random partitioning into training and validation sets with the training set used in rrBLUP to predict the phenotype of the validation set was iterated 10,000 times (utilizing all SNPs) with 10 repetitions at each

level for each trait from which the mean predictive accuracy (r) was determined. Correlations between observed and predicted values using the entire population (100%) in both the rrBLUP calculations and the cross validated rrBLUP calculations were determined in R using a Pearson correlation coefficient.

Associated SNPs from the GWAS analysis were not added to the rrBLUP model as fixed effects (Spindel et al., 2016) because of the relatively low R^2 values of variance explained by associated SNPs but we did investigate the effect of number of SNPs retained in the model on prediction accuracy. SNPs were sorted from lowest to highest P value. From these, nine subsets (in addition to the full set) were created. The full SNP sets had 7,082 for biomass and 8,032 for all other traits (number of SNPs retained after filtering for $MAF < 0.05$). These were reduced in an exponential manner (3,541, 1,770, 885, 442, 221, 120, 55, 28, and 14 SNPs for biomass and 4,018, 2,009, 1004, 502, 251, 126, 63, 32, and 16 for all other traits) to create the subsets. Each subset contained the most highly significant SNPs identified by GWAS. For each subset, the correlation of observed with predicted values was computed in rrBLUP.

RESULTS

ANOVA

Means and standard errors for the traits measured in the BeanCAP SBDP are shown in **Table 2**. Histograms (**Supplementary Figure 4**) based on LSMeans showed traits to be approximately normal in distribution except for biomass. Biomass was unimodal but right skewed for LSMeans. The lines making up the BeanCAP SBDP exhibited large differences for all of the traits evaluated. Mean squares for the ANOVA model were highly significant for all traits evaluated (**Table 3**). Mean squares for lines were either significant or highly significant for all traits

TABLE 2 | Means and standard error (SE) ($N = 16$), and narrow sense heritability (h^2) and 95% confidence intervals for heritability for *Fusarium solani* root rot symptoms (disease severity), plant biomass and root parameters of lines grown in the BeanCAP Snap Bean Diversity Panel at the Oregon State University Vegetable Research Farm in 2014 and 2015.

| Trait | Mean ¹ | SE (mean) | h^2 | 95% Confidence interval (h^2) |
|---------------------------|-------------------|-----------|-------|-----------------------------------|
| Disease severity | 3.10 | 0.01 | 0.74 | (0.66–0.82) |
| Aboveground Biomass | 3.28 | 0.02 | 0.75 | (0.67–0.83) |
| Adventitious Roots | 1.98 | 0.02 | 0.64 | (0.52–0.76) |
| Taproot Diameter (cm) | 2.27 | 0.02 | 0.51 | (0.35–0.67) |
| Basal Root Diameter (cm) | 2.08 | 0.02 | 0.47 | (0.29–0.64) |
| Shallowest Root Angle | 16.14 | 0.30 | 0.38 | (0.18–0.58) |
| Deepest Root Angle | 55.68 | 0.31 | 0.38 | (0.18–0.58) |
| Root Angle Average | 39.55 | 0.36 | 0.41 | (0.22–0.60) |
| Root Angle Difference | 35.91 | 0.25 | 0.32 | (0.10–0.54) |
| Root Angle Geometric Mean | 36.10 | 0.23 | 0.33 | (0.12–0.55) |

¹Disease severity rated on a 1–5 scale where 1 is resistant and 5 is susceptible; Biomass rated on a 1–5 scale where 1 is the least and 5 the most biomass accumulation; Adventitious roots rated on a 1–3 scale where 3 = most adventitious roots; and root angle measurements are in degrees from 0° to 90° where 0° represents a horizontal position.

¹<https://www.maizegenetics.net/tassel>

TABLE 3 | Degrees of freedom, mean squares, and significance level for model, year, bean line, replicate within year, and year by line interaction from an analysis of variance for traits associated with *Fusarium solani* disease reaction and plant and root parameters evaluated in trials at the Oregon State University Vegetable Research farm near Corvallis, of the BeanCAP Snap Bean Diversity Panel in 2014 and 2015.

| Source of variation | d.f. | Disease severity | Aboveground biomass | Adventitious roots | Taproot diameter | Basal root diameter |
|---|------|-----------------------|---------------------|-----------------------|--------------------|---------------------------|
| Disease, plant, and primary root traits | | | | | | |
| Model | 301 | 0.80*** | 2.40*** | 1.00*** | 0.77*** | 0.58*** |
| Year | 1 | 2.03 ^{ns} | 10.80 ^{ns} | 13.04 ^{ns} | 21.51* | 22.46* |
| Line | 147 | 1.25*** | 3.71*** | 1.33*** | 0.89*** | 0.62*** |
| Rep(Year) | 6 | 1.47*** | 4.97*** | 3.66*** | 2.58*** | 1.99*** |
| Year*Line | 147 | 0.32* | 0.92*** | 0.48*** | 0.43 ^{ns} | 0.33* |
| R ² | | 0.55 | 0.58 | 0.54 | 0.42 | 0.44 |
| CV | | 15.4 | 23.4 | 27.1 | 26.6 | 24.2 |
| Source of variation | d.f. | Shallowest root angle | Deepest root angle | Root angle difference | Root angle average | Root angle geometric mean |
| Derived root traits | | | | | | |
| Model | 301 | 186.4*** | 210.0*** | 240.0*** | 138.0*** | 103.8*** |
| Year | 1 | 5455.4* | 5101.0* | 6.2 ^{ns} | 5276.6* | 1242.2* |
| Line | 147 | 200.6* | 233.2* | 285.2* | 145.4*** | 118.6* |
| Rep(Year) | 6 | 494.0*** | 418.5* | 241.4 ^{ns} | 395.8*** | 155.7* |
| Year*Line | 147 | 123.8 ^{ns} | 145.0* | 196.2 ^{ns} | 85.2 ^{ns} | 79.2* |
| R ² | | 0.35 | 0.38 | 0.32 | 0.38 | 0.36 |
| CV | | 68.1 | 19.4 | 33.0 | 24.3 | 22.0 |

Shown at the bottom of the table are R² values and coefficient of variation values. R² is the regression coefficient for fit to the general linear model. ns = not significant; * = significant at $P < 0.05$; *** = significant at $P < 0.001$.

evaluated, with lower significance levels corresponding to the root angle measurements and the traits derived thereof. Mean squares for replicate were either significant or highly significant, except for the derived trait root angle difference. The mean square for year was significant for taproot diameter, basal root diameter, shallowest root angle, deepest root angle, root angle average, and root angle geometric mean. It was not significant for any other traits. In no cases were years highly significant. Year by line interaction was significant for disease, basal root diameter, deepest root angle, and root angle geometric mean. It was highly significant for aboveground biomass and adventitious roots (Table 3).

Multiple Correlation Analysis Among Traits

Disease severity was negatively correlated with aboveground biomass, basal root diameter, and taproot diameter (Table 4), and positively correlated with adventitious roots, shallowest root angle, and deepest root angle. Aboveground biomass, basal root diameter and taproot diameter were highly positively correlated (Table 4). Aboveground biomass and taproot diameter were negatively correlated with shallowest and deepest root angle. Basal root diameter showed the same negative relationship with shallowest root angle but did not have a significant correlation with deepest root angle. Shallowest and deepest root angles were positively correlated with each other.

Heritability

A range of heritabilities was observed for the different traits measured (Table 2). Aboveground biomass

and disease severity had the highest heritability with $h^2 = 0.75$ and 0.74 , respectively. The root angle traits had the lowest heritability, ranging from $h^2 = 0.32$ for root angle difference to $h^2 = 0.41$ for root angle average.

Genome-Wide Association Study

GWAS assumes normality (Goh and Yap, 2009). The disease severity and biomass datasets were normally distributed based on QQ plots of residuals generated from an ANOVA analysis of years, reps, and genotypes. The tap root diameter and basal root diameter datasets were also normally distributed for residuals after a square root transformation. Adventitious roots, short root angle, and deep root angle could not be made to conform to normality for their residuals. A GWAS analysis was conducted on all datasets, including square root transformed tap root diameter and basal root diameter. GWAS analysis of tap root diameter, basal root diameter, adventitious roots, short root angle, and deep root angle did not generate any significant SNP associations with a two PCA FarmCPU model.

A biplot of the first two PC axes (Supplementary Figure 3) revealed a clinal gradient along PCA 1 for center of domestication, with those lines clearly from the Mesoamerican center of domestication having strong positive scores and those with Andean background ranging from positive to negative scores. PCA 2 primarily separated European-bred small sieve cultivars from blue lake and pole bean types, but without discernable differentiation for Andean background cultivars. These results generally match our findings with Structure analysis (Wallace et al., 2018).

TABLE 4 | Pearson multiple correlation coefficients¹ for Least Square Means of the BeanCAP Snap Bean Diversity Panel evaluated for *Fusarium solani* disease and plant and root traits at the Oregon State University Vegetable Research Farm near Corvallis in 2014 and 2015.

| | Biomass | Adventitious roots | Basal root diameter | Taproot diameter | Shallowest root angle | Deepest root angle | Root angle difference | Root angle average | Root angle geometric mean |
|-----------------------|----------|--------------------|---------------------|------------------|-----------------------|---------------------|-----------------------|---------------------|---------------------------|
| Disease severity | −0.35*** | 0.26** | −0.27** | −0.40*** | 0.37*** | 0.32*** | −0.02 ^{ns} | 0.42*** | 0.19* |
| Biomass | | 0.13 ^{ns} | 0.21* | 0.19* | −0.37*** | −0.18* | 0.14 ^{ns} | −0.33*** | −0.04 ^{ns} |
| Adventitious roots | | | −0.09 ^{ns} | −0.24** | 0.03 ^{ns} | 0.13 ^{ns} | 0.09 ^{ns} | 0.10 ^{ns} | 0.13 ^{ns} |
| Basal root diameter | | | | 0.37*** | −0.17* | −0.05 ^{ns} | 0.10 ^{ns} | −0.13 ^{ns} | 0.03 ^{ns} |
| Taproot diameter | | | | | −0.23** | −0.46*** | −0.23** | −0.43*** | −0.39*** |
| Shallowest root angle | | | | | | 0.34*** | −0.53*** | 0.80*** | 0.00 ^{ns} |
| Deepest root angle | | | | | | | 0.62*** | 0.83*** | 0.93*** |
| Root angle difference | | | | | | | | 0.08 ^{ns} | 0.84*** |
| Root angle average | | | | | | | | | 0.59*** |

¹Probability > |r| under H_0 : $\rho = 0$. * = significant at $P < 0.05$; ** = significant at $P < 0.01$; and *** = significant at $P < 0.0001$. ^{ns} = not significant.

Five SNPs were associated with disease severity on chromosomes Pv03, Pv07, Pv08, and Pv10 with two SNPs on Pv10 (Table 5 and Figure 1). SNPs ss715639797, ss715649485, and ss715646318 on Pv08 and Pv10 were identified through a Bonferroni threshold. A further two SNPs, ss715647578 and ss715646526, were identified on Pv03 and Pv07 through a bootstrap analysis. The phenotypic variation (R^2) explained by SNPs indicated a low contribution to disease resistance by each SNP ranging in value from 0.9 to 10.8% with the highest value for ss715647578 on Pv03 and the lowest value for ss715646526 on Pv07. The effect of allelic substitution was negative for three SNPs and positive for two (Table 5). Effect was relatively small with a cumulative effect of altering disease severity score by 0.5. Two SNPs were associated with biomass on chromosomes Pv10 and Pv11 (Table 5 and Figure 1). SNPs ss715649390 and ss715645486 on Pv10 and Pv11, respectively, were identified through a Bonferroni threshold. No further SNPs were identified through a bootstrap analysis. The R^2 values were 11.3% for ss715645486 and 14.8% for ss715649390, and the former had an allelic substitution effect of -0.12 while the latter had a relatively larger effect of -0.18 (Table 5). The cumulative effect of these two SNPs would be to shift the five-point scale by 0.3.

Within a 100 kb window upstream and downstream of these SNPs, a total of 123 gene models were found across the seven regions with an average of 18 per region (Table 5). One candidate gene (peroxidase) was identified as potentially involved in disease resistance (Table 6). A total of four candidate genes were identified as potentially involved in biomass and abiotic stress tolerance, including a pentatricopeptide repeat domain and three tandem 2-alkenal reductase genes models (Table 6). Two of the three 2-alkenal reductase gene models were outside of the 100 kb window, but are included here because they were adjacent to one within the window.

Based on a threshold of D' or $R^2 \geq 0.80$ and $P \leq 0.01$, regions of LD were identified around some significant SNPs. D' identified extremely large blocks of LD that were on the order of 1.9–36.0 Mb for disease severity whereas R^2 provided a much more conservative estimate, ranging from 150 to 679 kb (Supplementary Table 2). The LD heat map and table indicated

that SNPs on Pv03, Pv07, and Pv10 for disease severity, and Pv10 for biomass were within blocks of LD (Figure 2). These ranged from 150 to 679 kb in size. The other SNPs were in LD blocks using D' as a criterion, but not with R^2 (Supplementary Table 2).

An ANOVA analysis of the trait-SNP associations supported the results of the GWAS analysis (Supplementary Figure 5). The results were uniform for years with no significant differences between years. The box plot trend supported the trait-SNP association for SNP ss715639797 with $P = 0.08$. The other SNPs for disease were significant with $P < 0.05$. The only exception was SNP ss715646526 which was not significant, and the box plots did not show any particular trend, and this was true for individual years. For biomass, ss715649390 was highly significant whereas ss715645486 was not, but it does show a trend.

Genomic Prediction

GEBV rankings represent the general trends seen in the phenotypic data but with numerous crossovers in ranking due to the information from relatives reflected in GEBV calculations (Table 7). This can be seen in the ranking of disease severity, which has 'Impact', 'Black Valentine', 'Widusa', 'NY6020-5', and 'Romano Gold' as the top five most resistant lines in the phenotypic data set (data not shown) but the GEBV calculations show 'Widusa', 'Impact', 'Double Dutch White', 'Booster', and 'Stringless French Filet' as having the best GEBV for disease resistance (Table 7). When compared to the PCA biplot (Supplementary Figure 3), lines with the highest GEBV rankings for disease severity come from both Mesoamerican and Andean centers and provide evidence that population structure is not influencing choice of significant SNP associations.

Predicted and observed values for all traits resulted in high to moderate correlations (r) for disease severity, biomass, and the five root architecture traits (Figure 3 and Supplementary Table 3; 100% column in histograms and row in table). Ten thousand iterations of a cross validation with four training-testing models and replicated 10 times for each trait-model combination produced moderate to low correlations for predictive ability. The correlations that were highest under training and validation were those for disease severity, biomass

TABLE 5 | SS identification numbers of the SNP, chromosome, position, negative log p -value, minor allele frequency (MAF), proportion of total phenotypic variation explained by the SNP (R^2), allelic effect, chromosomal location and number of gene models found within a 200 kb window proximal and distal to the SNP for significant associations found from genome wide association study of *Fusarium solani* root rot disease severity and biomass in the BeanCAP Snap Bean Diversity Panel grown at the Oregon State University Vegetable Research Farm in 2014 and 2015.

| Trait | SS ID No. | Chromosome | Position (bp) | $-\log P$ | MAF | R^2 | Effect | Chromosomal location ¹ | No. gene models |
|---------|-------------|------------|---------------|-----------|------|-------|--------|-----------------------------------|-----------------|
| Disease | ss715647578 | Pv03 | 12,661,037 | 4.58 | 0.09 | 10.8 | −0.15 | pericentric | 11 |
| Disease | ss715646526 | Pv07 | 34,296,485 | 4.51 | 0.37 | 0.9 | 0.09 | pericentric | 21 |
| Disease | ss715639797 | Pv08 | 32,951,182 | 5.24 | 0.23 | 6.2 | −0.13 | pericentric | 7 |
| Biomass | ss715649390 | Pv10 | 5,677,538 | 4.89 | 0.37 | 14.8 | −0.12 | proximal | 11 |
| Disease | ss715649485 | Pv10 | 7,910,750 | 4.85 | 0.14 | 7.3 | −0.13 | pericentric | 12 |
| Disease | ss715646318 | Pv10 | 40,686,027 | 5.75 | 0.39 | 5.6 | 0.14 | distal | 35 |
| Biomass | ss715645486 | Pv11 | 766,814 | 5.35 | 0.22 | 11.3 | −0.18 | proximal | 26 |

¹ Pericentric location of a SNP is associated with low rates of recombination while proximal and distal locations are in regions of high recombination. Placement based on **Supplementary Figures** of physical vs. linkage map distances in Schmutz et al. (2014).

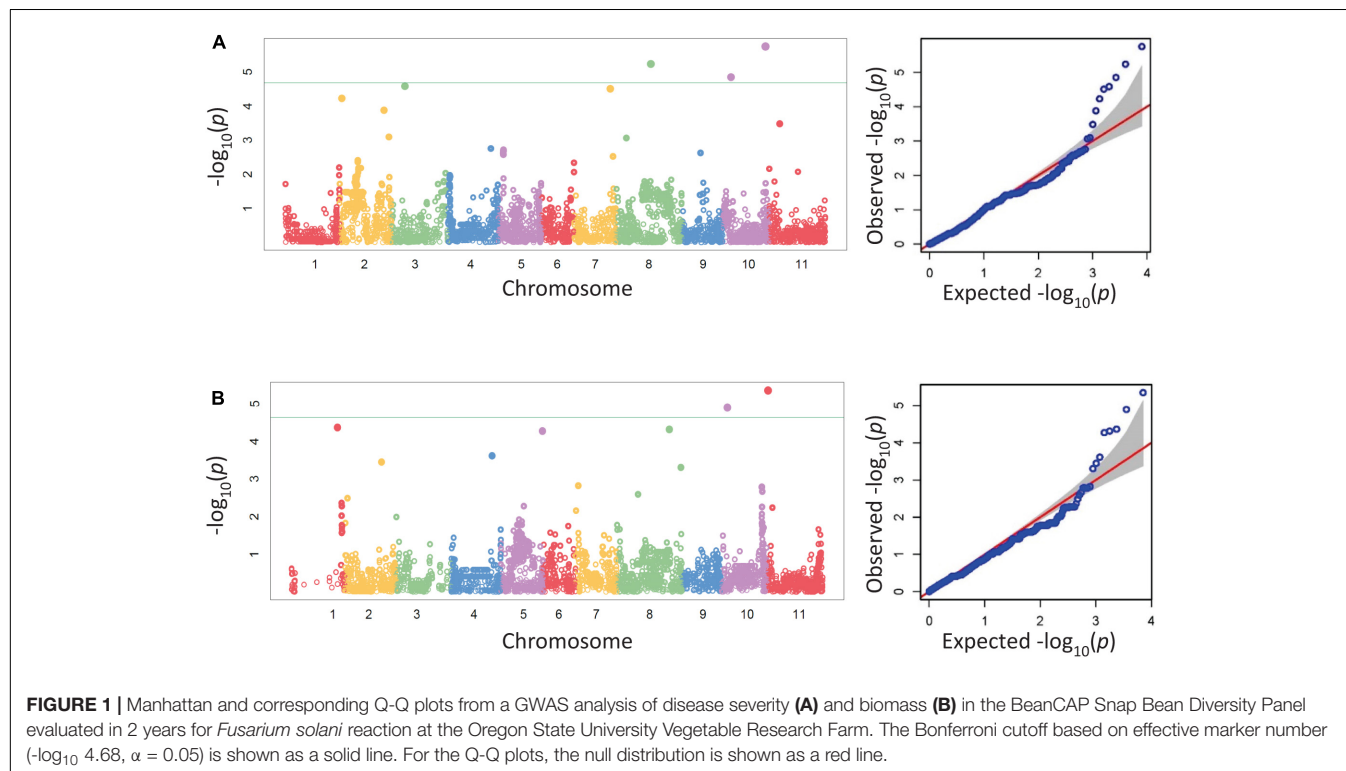


TABLE 6 | Putative candidate genes within 350 kb of the associated SNP for *Fusarium solani* root rot disease severity and plant biomass identified by genome wide association study using the BeanCAP snap bean diversity panel grown at the Oregon State University Vegetable Research Farm near Corvallis.

| Chrom. | SNP position | Distance ¹ | <i>P. vulgaris</i> gene model | Start | End | Gene function | References |
|-------------------|--------------|-----------------------|-------------------------------|------------|------------|---|--------------------------------------|
| | | bp | | | bp | | |
| Pv03 ² | 12,661,037 | 74,361 | Phvul.003G078600.1 | 12,584,313 | 12,586,676 | Peroxidase | Ray et al., 1998 |
| Pv10 | 5,677,538 | 76,042 | Phvul.010G039100 | 5,753,580 | 5,758,499 | 2-alkenal reductase | Xi et al., 2015 |
| | | 102,406 | Phvul.010G039200 | 5,779,944 | 5,784,500 | 2-alkenal reductase | Xi et al., 2015 |
| | | 125,106 | Phvul.010G039300 | 5,802,644 | 5,808,058 | 2-alkenal reductase | Xi et al., 2015 |
| Pv11 | 766,814 | 72,566 | Phvul.011G010900 | 839,380 | 841,056 | Pentatricopeptide repeat domain (PPR_3) | Jiang et al., 2015; Cao et al., 2020 |

¹ Distance between SNP and nearest end of candidate gene. ² QTN on Pv03 associated with disease severity while those on Pv10 and Pv11 are associated with biomass.

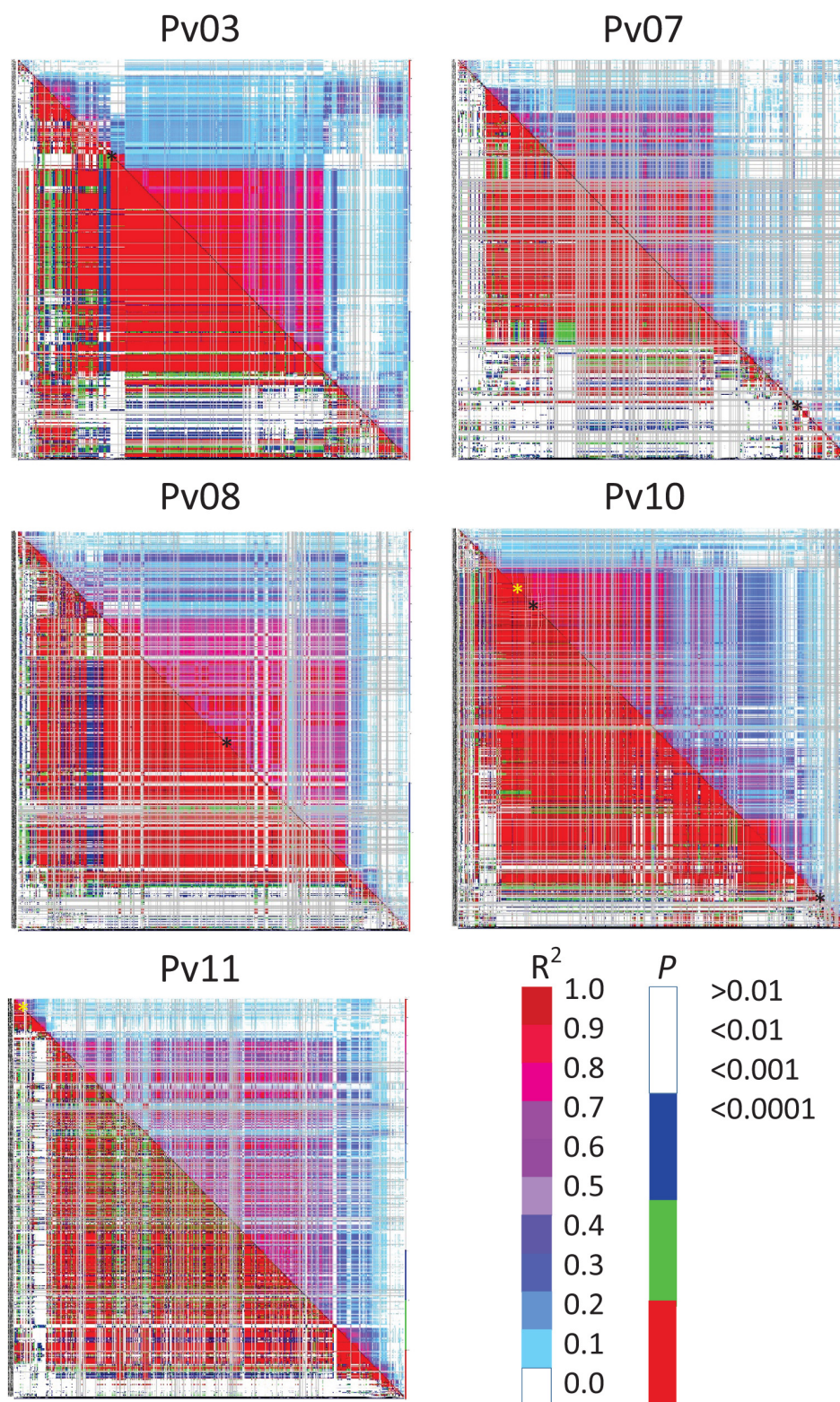


FIGURE 2 | Linkage disequilibrium (LD) heat map of common bean chromosomes Pv03, 07, 08, 10, and 11 showing all possible pairwise comparisons of SNPs arranged along the chromosome. R^2 values are displayed above and right of the diagonal and corresponding probabilities below and left of the diagonal. Color scales show corresponding R^2 and probabilities where red for each would indicate strong and highly significant LD. SNPs associated with disease severity (black *) or biomass (yellow *) are indicated along the diagonal.

TABLE 7 | Genomic estimated breeding values (GEBV) calculated from BLUPs and BLUEs for the 10 highest and 10 lowest ranked lines in the BeanCAP Snap Bean Diversity Panel for *Fusarium solani* root rot disease severity, plant biomass, tap root, basal root diameter, adventitious roots, deepest and shallowest root angle.

| Disease severity | | Biomass | | Tap root diameter | | Basal root diameter | | Adventitious roots | | Deep root angle | | Shallow root angle | |
|-------------------------|-------------------|-----------------|-------------------|-------------------|-----------|-------------------------|-----------|--------------------|-------------------|-----------------|-------------------|--------------------|-------------------|
| Accession | GEBV ¹ | Accession | GEBV ² | Accession | GEBV (mm) | Accession | GEBV (mm) | Accession | GEBV ³ | Accession | GEBV ⁴ | Accession | GEBV ⁴ |
| Widusa | 2.61 | Oregon 2065 | 4.32 | Widusa | 2.71 | Goldrush | 2.28 | Widusa | 1.7 | Booster | 48.8 | Oregon Giant Pole | 11.51 |
| Impact | 2.64 | Idaho Refugee | 4.15 | Trail of Tears | 2.66 | Oregon 91G | 2.27 | Serin | 1.71 | Oregon 2065 | 48.88 | Roma II | 11.68 |
| Dutch Double White | 2.68 | Corbett Refugee | 4.04 | Fortex | 2.65 | Gold Mine | 2.26 | Pole Blue Lake S7 | 1.72 | Banga | 49.08 | Fortex | 11.99 |
| Booster | 2.72 | Gina | 4 | Pole Blue Lake S7 | 2.62 | Profit | 2.26 | Dutch Double White | 1.74 | Pole Blue Lake | 49.09 | Ebro | 12.02 |
| Stringless French Filet | 2.73 | Ebro | 3.97 | Booster | 2.56 | Stringless French Filet | 2.24 | Impact | 1.74 | Serin | 49.26 | Magnum | 12.05 |
| Selecta | 2.74 | Tapia | 3.92 | EZ Pick | 2.54 | Oregon 5630 | 2.22 | Kylian | 1.76 | Astun | 49.3 | Tapia | 12.11 |
| Pole Blue Lake | 2.78 | NY6020-5 | 3.89 | Impact | 2.54 | Eagle | 2.22 | Koala | 1.76 | EZ pick | 49.36 | Astun | 12.6 |
| Oregon 2065 | 2.79 | Coloma | 3.89 | Pole Blue Lake | 2.53 | Gina | 2.22 | Polder | 1.77 | Celtic | 49.44 | Idaho Refugee | 12.96 |
| Pole Blue Lake S7 | 2.79 | Unidor | 3.87 | Paloma | 2.51 | Carson | 2.21 | Renegade | 1.77 | Stayton | 49.46 | Romano 118 | 13.01 |
| Cherokee | 2.81 | Calgreen | 3.82 | Hayden | 2.5 | Summit | 2.21 | Pix | 1.78 | Redon | 49.67 | Cyclone | 13.01 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Shade | 3.45 | Brio | 2.68 | US Refugee #5 | 2.05 | Redon | 1.96 | NY6020-5 | 2.2 | Benton | 59.87 | Benton | 18.51 |
| Espada | 3.45 | Minuette | 2.67 | Charon | 2.04 | EZ Pick | 1.94 | Medinah | 2.2 | Castano | 59.89 | Warrior | 18.66 |
| Spartacus | 3.46 | Paulista | 2.66 | Opus | 2.04 | Banga | 1.93 | Landmark | 2.2 | Brio | 59.91 | Festina | 18.68 |
| Matador | 3.49 | Festina | 2.65 | Strike | 2.04 | Idaho Refugee | 1.93 | Benton | 2.24 | Shade | 60.04 | Zeus | 18.68 |
| Warrior | 3.5 | Matador | 2.64 | Mercury | 2.04 | Booster | 1.93 | Coloma | 2.29 | Summit | 60.1 | Matador | 18.72 |
| Titan | 3.53 | Palati | 2.64 | Dusky | 2.03 | Blue Peter Pole | 1.91 | FR-266 | 2.32 | Carlo | 60.38 | Palati | 18.77 |
| Benton | 3.53 | Flavorsweet | 2.63 | Castano | 2.02 | Corbett Refugee | 1.89 | Oregon Giant Pole | 2.34 | Provider | 60.56 | Benchmark | 18.88 |
| Hercules | 3.53 | Dusky | 2.51 | Landmark | 2.01 | Kentucky Wonder | 1.87 | US Refugee #5 | 2.43 | Stallion | 60.82 | Dusky | 19.04 |
| Festina | 3.54 | Speedy | 2.4 | Idaho Refugee | 1.97 | McCaslan No. 42 | 1.86 | Idaho Refugee | 2.55 | Valentino | 61.39 | Castano | 19.22 |
| Seabiscuit | 3.58 | Embassy | 2.32 | Corbett Refugee | 1.93 | Trail of Tears | 1.86 | Corbett Refugee | 2.6 | Grenoble | 61.41 | Roller | 19.37 |

¹Disease severity rated on a 1–5 scale where 1 is resistant and 5 is susceptible. ²Rated on a 1–5 scale where 1 is the least and 5 the most biomass accumulation. ³Rated on a 1–3 scale where 3 = most adventitious roots.

⁴Degrees from 0° to 90° where 0° represents a horizontal position.

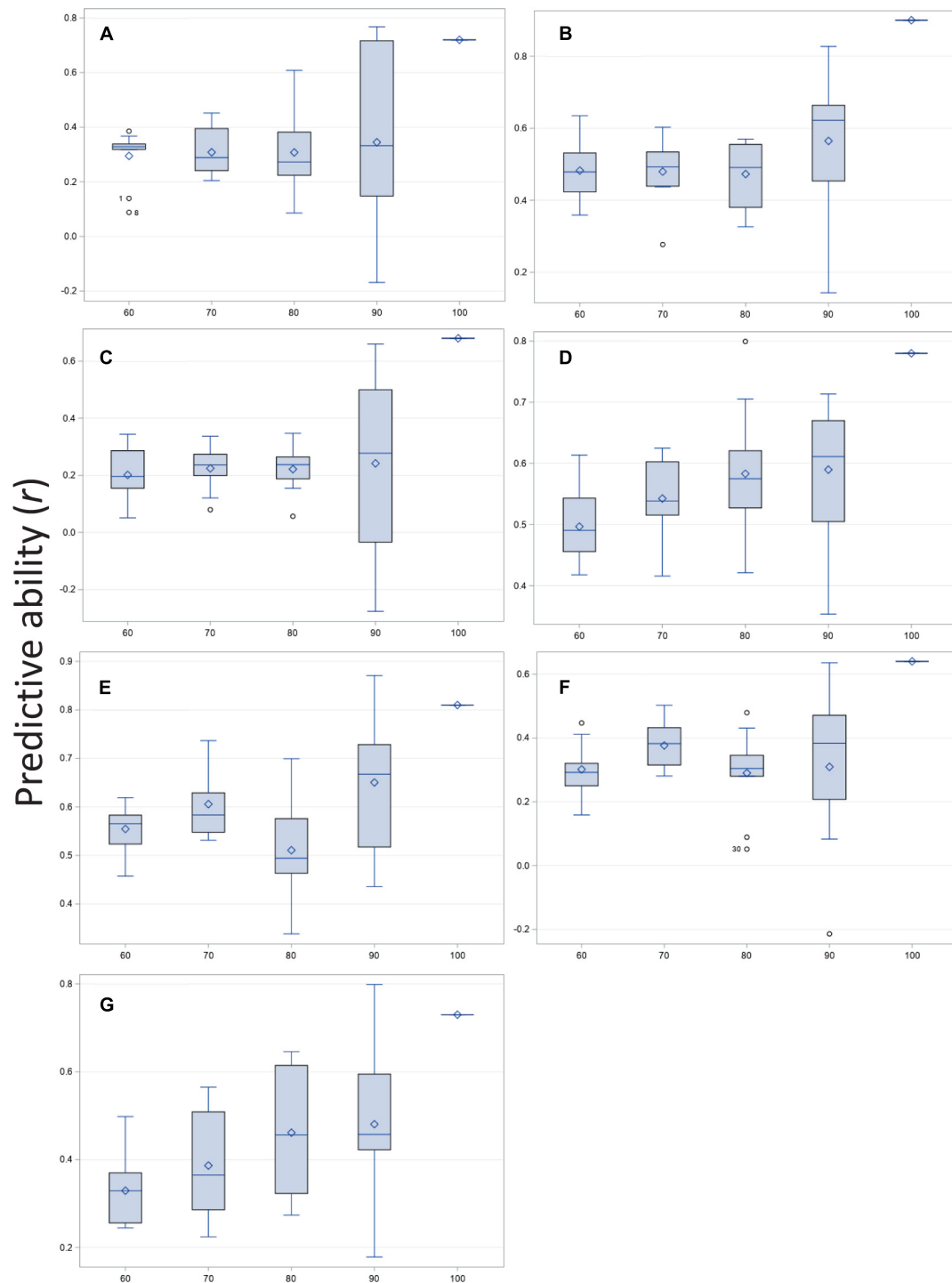


FIGURE 3 | Correlation of predicted and observed values for training vs. testing populations at four ratios (60:40, 70:30, 80:20, and 90:10%) and compared with observed vs. predicted for the entire population (100%) of the BeanCAP Snap Bean Diversity Panel for *Fusarium solani* root rot disease severity and plant and root traits. Correlation coefficients were generated by rBLUP using 10 K iterations and 10 repetitions per trait-level combination. **(A)** Adventitious roots, **(B)** Biomass, **(C)** Basal stem diameter, **(D)** Disease severity, **(E)** Deep root angle, **(F)** Shallow root angle, and **(G)** Taproot diameter.

and deep root angle. As size of the training population increased, mean correlation remained essentially flat (adventitious roots, basal root diameter), showed a linear increase (biomass, disease

severity, and taproot diameter), or fluctuated (deep and shallow root angles). Variation about the mean of r was greatest at the 90% level (Figure 3 and Supplementary Table 3). Overall,

standard deviations were smallest for the model with 70% training population although for biomass, 60 or 70% training models were very similar, as were 70 and 80% training models for basal root diameter. Cross-validation predictions generally were 20–40% lower than correlation among predicted and observed of the entire population. Disease severity, deep root angle and shallow root angle showed the smallest differences.

Number of SNPs retained in the model affected predictive ability. Correlation coefficients were generally lowest for the fewest significant SNPs and increased as SNPs were added to the model (Figure 4), but in most cases plateaued before declining with use of the full SNP set. The traits separated into two groups with disease severity and biomass showing relatively high correlations, and the root traits exhibiting moderate to moderately high correlations over SNP subsets. For disease severity, $r > 0.90$ was obtained with 126 SNPs, while for biomass $r > 0.90$ was obtained with 221 SNPs. Disease severity exhibited a decrease in r from 0.91 to 0.78 when transitioning from 4,018 to the full SNP set, and for biomass, the decrease was from 0.93 to 0.90. For root traits, most did not reach a maximum r until 2,009 or 4,018 SNPs were used with r ranging from 0.72 to 0.84. In all cases except for adventitious roots and deep root angle, r decreased for the full SNP set compared to half the SNPs used in the model (Figure 4).

DISCUSSION

Our ANOVA results showed significant year \times line interaction for disease severity ($P \leq 0.05$), biomass ($P \leq 0.001$), adventitious

roots ($P \leq 0.001$), basal root diameter ($P \leq 0.05$), deepest root angle ($P \leq 0.05$), and root angle geometric mean ($P \leq 0.05$), but no statistical significance for the remaining traits. The significant interactions for disease severity and biomass appeared to be due to differences in magnitude rather than changes in rank years as shown by moderate but highly significant correlations between years based on Spearman rank correlation (data not shown). The pattern exhibited by the replicates for disease score differed in 2014 and 2015, most likely due to differences in order of evaluation. In 2014, lines in all reps were evaluated when reaching the desired physiological stage but in 2015, reps were evaluated sequentially. In 2014, spatial variation in reps was important with the two inner reps showing more disease than the outer reps. In 2015, disease severity increased over time. Coefficient of variation (CV) was relatively low at 15 and 23 for disease severity and biomass, respectively, with other traits similar to biomass, except shallow root angle, which was had a CV of 68. The high disease pressure and consistent watering likely contributed to this uniformity across years and a low CV. Although our study could not exclude every environmental factor present in an outdoor field, these environmental factors may be both confounding but also offer the possibility of capturing complex interactions between genes and the environment that could be important to disease manifestation in a grower's field.

Our shovelomics methodology provides a valuable window into the disease process. Our analysis showed that root angle and disease severity are positively correlated suggesting that susceptible lines had root systems oriented at a deeper angle than resistant lines (Table 4). Similar to our findings, Snapp et al. (2003) found that more lateral roots of larger diameter were

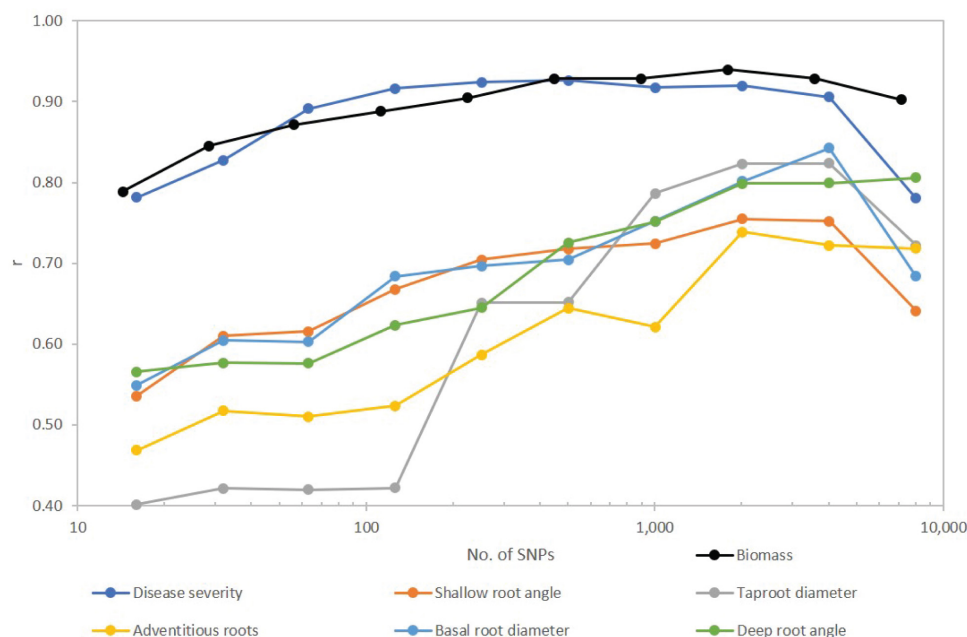


FIGURE 4 | Effect of number of SNPs on predictive accuracy for *Fusarium* root rot disease severity and root traits of a snap bean diversity panel. SNPs were first filtered for MAF < 0.05 , then sorted from smallest to largest P value and arranged in nine subsets approximately doubling in size with each step. Full set of SNPs for biomass was 7,082 while for all other traits totaled 8,036. Number of SNPs is plotted on a logarithmic (base 10) scale.

associated with *Fusarium* root rot resistance. In their research on nutrient foraging, Lynch and Brown (2001) emphasized that a plant with exclusively deep root angles is exploring a smaller amount of soil than a plant with either a shallow or a range of root angles. The beans with shallower root systems may have been able to access a greater soil volume. Another possible explanation for the effect observed in this study is that the upper layer of soil had superior drainage, which reduced infection by root rot. There may be a tradeoff between disease resistance and drought tolerance with regard to root angle. Drought tolerant plants will likely have roots exploring greater depths of soil.

The negative correlation of disease severity with aboveground biomass, basal root diameter, and taproot diameter, indicated that resistant cultivars had greater aboveground biomass and larger root diameter than susceptible cultivars (Table 4). The positive correlation of disease severity with adventitious roots, shallowest root angle and deepest root angle indicated that cultivars with fewer adventitious roots and shallower root angles were associated with less disease. For aboveground biomass, basal root diameter and taproot diameter, the highly positive correlation indicated that the magnitude of the three size measurements maintained a constant relationship across lines. Aboveground biomass and taproot diameter were negatively correlated with shallowest and deepest root angle, meaning that larger plants had shallower root systems. Positively correlated shallowest and deepest root angles indicated that regardless of the orientation of the root system, the span of the soil profile that it accessed remained constant.

Disease severity, biomass, and adventitious roots had higher heritability than the other shovelomics traits, such as root angle measurements. The heritability value for disease resistance is within the range of values measured by most previous researchers. Hagerty et al. (2015) obtained h^2 of ~ 0.20 and Mukankusi et al. (2011) reported heritability of 0.38–0.45 for root rot resistance. In contrast, Kamfwa et al. (2013) found higher heritabilities of 0.86–0.99. The heritability for aboveground biomass found in this study also corresponds to previously reported values. Shenkut and Brick (2003) found a range from 0.60 to 0.70. Navarro et al. (2008) reported values of 0.77–0.91 for heritability of biomass, based on measurements of dry weight, which implies that our categorical rating system did not greatly inflate heritability values. The high heritability values imply that simple selection strategies on these traits would be effective.

The high heritabilities of disease severity and biomass are consistent with the high correlational accuracy of these two traits in genomic prediction and the significant results in GWAS. These two traits were also negatively correlated with a high statistical significance (Table 4) indicating the possibility that disease stressed plants were generating less biomass. Nevertheless, these traits are not entirely overlapping and the negative correlation may be partly coincidental and not causal because GWAS analysis identified distinct SNP markers for disease severity and biomass.

The lack of GWAS results for five of seven traits is notable. There may be confounding factors associated with measuring traits under disease pressure. As noted already, the other traits had lower heritabilities that may also explain the difference. Moreover, the Bonferroni and bootstrap thresholds utilized

in this study are very conservative. Additionally, increasing the population size and/or number of SNPs would have led to greater precision and a greater likelihood of detecting significant associations.

The SNPs identified by our GWAS analysis did not clearly overlap with any previously identified SNP from GWAS analysis or biparental analysis of root rot organisms (Hagerty et al., 2015; Oladzad et al., 2019; Dramadri et al., 2020; Zitnick-Anderson et al., 2020). We identified one candidate gene related to plant defense within the immediate vicinity of an associated SNP (Table 6). Peroxidases are involved in the final steps of the biochemical pathway leading to lignification, which directly interferes with pathogen invasion (Ray et al., 1998).

From our studies and those of others (Hagerty et al., 2015; Nakedde et al., 2016; Wang et al., 2018; Zitnick-Anderson et al., 2020), there is strong evidence that *F. solani* resistance in common bean is polygenic with many genes with small effect being involved. One interesting finding is the lack of commonality of resistance QTL among the different studies where genome location can be compared. This would support the idea of polygenic resistance based on genes that are not considered classical resistance genes. Given the level of resistance in some lines in our diversity panel, it is possible to achieve relatively high levels of resistance with the right gene combination, which appears to confer broad-spectrum resistance to different *Fusarium* isolates. While virulence may vary among isolates, there does not appear to be a pathogen race structure. As a case in point, the resistance in FR-266 was relatively effective to *Fusarium* isolates endemic to Michigan (Schneider et al., 2001; Snapp et al., 2003), whereas we found this genotype to be moderately resistant against our field isolates in Oregon, implying that Oregon isolates were more virulent. However, in both cases, resistance was quantitative with no clear major QTL.

Where host and pathogen are coevolving under antagonist selection, the prediction is resistance genes would evolve in concert and tend to accumulate in large haplotype blocks in low recombining genomic regions (Ravinet et al., 2017). Our findings lend support to that idea in that of the five SNPs associated with disease severity, four were located in low-recombination, gene-sparse pericentric regions and only one was located distally on Pv10 in a high-recombination region (Table 5). Both SNPs associated with biomass were in high-recombination regions located proximally on their respective chromosomes.

Linkage disequilibrium heatmaps (Figure 2 and Supplementary Table 2) provide a more detailed examination of low recombination blocks in relations to chromosomal location, and are in partial agreement with low recombination regions identified in Table 5. Visually, Figure 2 aligns with categories in Table 5. One discrepancy between Table 5 and Supplementary Table 2 was for the SNP associated with disease severity on Pv08, where the SNP clearly resides in a region of low recombination (based on physical vs. cM biplots in Schmutz et al., 2014), however, an LD block for this region was essentially non-existent based on an R^2 cutoff of 0.80. The heatmap (Figure 2) does show moderate to high LD in this region. The second discrepancy was for a SNP on Pv10 associated with biomass. This SNP is located proximally, but had a sizable LD block of 421 kb. Pv10

is acrocentric and the SNP is located in the short arm, which have reduced recombination (see **Supplementary Figure 13** in Schmutz et al., 2014).

A further implication of the location of most resistance associated SNPs in low recombination regions is that marker assisted selection would be at best, inefficient and at worst, ineffective because of the large non-recombinant blocks of genes. This provides further support for prioritizing genomic selection over QTL mapping and marker assisted selection of individual QTL.

The biomass candidate genes were identified through their known effects on biomass but also their effects on abiotic stress tolerance because disease pressure can induce drought stress in affected bean plants through the loss of their roots to disease. A pentatricopeptide repeat (PPR) domain candidate gene was found in the vicinity of SNP ss715645486. PPR domains have been implicated in an increase of biomass in a study of *Paulownia* trees (Cao et al., 2020), and are also implicated in drought stress tolerance (Jiang et al., 2015). The three tandem duplicate genes of 2-alkenal reductase in the vicinity of SNP ss715649390 are also implicated in increased biomass and improved drought tolerance in a study of transgenic tobacco plants (Xi et al., 2015).

Are there tradeoffs between *Fusarium* resistance and abiotic stress tolerances? Burke and Miller (1983) extensively analyzed the interactions of *Fusarium* root rot with various cultural practices that can affect the development of disease. Their findings were that anything that constricts the root system (such as cold soils and compaction) will exacerbate disease development. Intermittent drought stress combined with these factors restricting root growth will further increase disease pressure. Excess soil moisture even if it is intermittent and of short duration will prevent oxygen diffusion to the roots and further inhibits root growth. High population densities also tend to increase root rot. Previously bred *Fusarium* root rot resistant dry bean cultivars tended to tolerate cold soils, drought and compaction better than susceptible cultivars, but in waterlogged soils, resistance was defeated. In the present study, there does not appear to be a tradeoff among these traits with one exception: the correlation of shallow root angle with disease resistance, which might lead to less drought tolerant plants. Correlation is not causation so this supposition would need to be tested and could be carried out by subjecting the snap bean diversity panel to drought as well as other forms of abiotic stress. On the other hand, nutrient use efficiency, especially for phosphorous (P), is associated with shallow root systems (Lynch and Brown, 2001). Breeding for P use efficiency would not likely impact root rot resistance and vice versa.

The multiple associated SNPs detected for disease severity with low R^2 values and their non-overlap with numerous SNPs detected for root rot in other studies strongly suggests that root rot resistance is highly polygenic in nature with numerous loci of low effect. This further supports the notion that genomic selection, which fully utilizes all SNPs, may be a better method to breed for root rot disease resistance in snap bean than identifying a small number of loci in GWAS and applying marker assisted selection to those loci.

Optimum ratio of training to testing populations for achieving the highest repeatable predictive ability was 70:30%

training:validation for most traits. This level is within the range of what has been found for other studies of genomic prediction in common bean (Keller et al., 2020; Diaz L. M. et al., 2021; Diaz S. et al., 2021; Shi et al., 2021). At 90% training population, the highest average predictabilities as measured by r were achieved, but standard deviations were much larger, leading to less certainty in whether a prediction was accurate. Shi et al. (2021) reported that training sets >80% can lead to large variation associated with too small a validation set.

In evaluating the influence of the number of SNPs on prediction accuracy, it was curious that for most traits, the full set of SNPs used in our model had lower predictive accuracy compared to a reduced number of SNPs. Studies in bean and other crops have generally shown a positive correlation between number of SNPs and predictive accuracy (Spindel et al., 2016; Liu et al., 2018; Wen et al., 2019; Keller et al., 2020; Thistlethwaite et al., 2020; Arenas et al., 2021; Shi et al., 2021). These studies do differ in how many SNPs were used and in how they were selected for each subset, but the overall trends were similar. Some studies have observed decreases in predictive accuracy at various SNP levels. Thistlethwaite et al. (2020) observed a drop at around 10,000 SNPs before rising again. Arenas et al. (2021) observed a dip at around 1,000–1,500 SNPs for four traits. In our study, disease severity and biomass could be modeled with a high degree of accuracy ($r > 0.90$) with relatively few (126–221) SNPs. In contrast, root traits were best modeled with one-half to one-quarter of the full SNP data set. Other studies have shown that genomic prediction models that incorporate GWAS can improve accuracy in breeding programs (Spindel et al., 2016). Shi et al. (2021) found the highest predictive accuracy when 20 SNPs derived from GWAS were used. Our selection of 14 (biomass) and 16 (disease severity) most highly significant SNPs had among the lowest predictive accuracies. Our results reinforce the idea that resistance to *Fusarium* root rot is polygenic and requires many genes to achieve the highest levels of resistance.

One of the important questions in GWAS has been how to account for the “missing heritability” in such studies (Manolio et al., 2009). Relative to the heritability estimates based on phenotypic and genotypic variances, the amount of variation explained by significant SNP associations is small, and the cumulative effect of all associations in the model does not always approximate classical measures of heritability. This is particularly true where QTL have small individual effect. In the present study, the h^2 estimate based on genotypic and phenotypic variances was relatively high 0.74 for disease severity and 0.75 for biomass (Table 2) while the cumulative R^2 for the SNPs associated with these traits ranged from 0.26 to 0.31. This implies that either h^2 is overestimated, or that GWAS may be missing medium- and low-effect associations. Relaxing our cutoff for identifying SNP associations could lead to the identification of additional associations, but increasing number of genotypes and/or markers would provide the greatest possibility of accurately detecting additional associations.

One piece of the missing heritability may be conditioned by genetic variability in the phenolic/flavonoid biosynthetic pathway. Flavonoids and phenolics have been shown to possess antimicrobial properties which have been associated with resistance to root rots

(Hagerty et al., 2015; Cirak and Myers, 2021). One line ('Cherokee') from those with the highest rank for GEBV for disease severity had colored seeds and flowers, while none of the lowest ranked lines were colored (Table 7 and Supplementary Table 1 in Kleintop et al., 2016). The SBDP has been evaluated for total phenolic content (TPC) of pods (Kleintop et al., 2016), which can serve as a proxy for phenolics and flavonoids distributed in other plant parts. The 10 lines with lowest GEBV values for disease severity had relatively higher TPC than did the 10 lines with the highest GEBV (mean of 0.52 vs. 0.40 mg g⁻¹ FW gallic acid equivalents). Disease severity and GEBV for disease severity were negatively correlated with TPC ($r = -0.18$, $P = 0.03$ and $r = -0.23$, $P = 0.005$, respectively). Myers et al. (2019) conducted a GWAS for TPC in pods of the SBDP and when we compared those results to the current findings, we did not find any overlap in regions of significant SNPs for disease severity or biomass. These results are compatible with the idea that phenolics do play a role in root rot resistance although it is not a major one.

To achieve acceptable processing quality, most contemporary snap bean cultivars are white-seeded, which eliminates anthocyanins and flavonols from the pods. If we had found a strong relationship between TPC and disease severity, those associations with pigment production would not be useful in a breeding program. Although lines varied for total TPC, all but one was white-seeded (preventing anthocyanin accumulation in the pods) and thus do not present barriers to use in a breeding program for root rot resistance.

In common bean, geographic origin and population structure have been shown to be an important influence on genetic variation in wild and landrace beans (Blair et al., 2012; Cortés et al., 2018). With the BeanCAP snap bean diversity panel, we do not expect associations that might be related to demography since snap bean origins are not associated with a particular place. However, snap beans do appear to have been secondarily derived from dry beans, and indirectly from the two centers of domestication, possibly with several independent events, and have retained some genetic signature of their origins (Wallace et al., 2018). Derivation has been followed by substantial admixing, which has reduced distinct associations with centers of domestication and has produced more of a clinal variation across the diversity panel. Population structure could result in spurious marker – trait associations; however, structure was accounted for in the FarmCPU model, and we did not see any pattern between disease severity GEBVs and location on the PCA biplot.

This research builds on prior work on *Fusarium* root rot resistance in common bean and will give snap bean breeders additional tools to dissect and manipulate resistance to *Fusarium* root rot in snap beans. The heritabilities give information on the expected gain from selection that could be achieved. The correlations among disease and root traits provide valuable information on the root architecture necessary to develop resistant lines. The GWAS analysis provides additional markers to a growing number associated with resistance. The genomic predictions identify individual lines with genetic merit worth pursuing by utilizing the totality of marker effects. Future research could include a more detailed investigation root trait associations with biotic and abiotic stress tolerance,

combine snap bean data with dry bean for a meta-GWAS, and development of a MAGIC population (Cavanagh et al., 2008) to facilitate recombination of SNP associations into a common snap bean background.

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: ScholarsArchive@OSU: <https://ir.library.oregonstate.edu/concern/datasets/m900p1589> for SNP data; Bean CAP Snap Bean Diversity Panel passport data are also available: <https://ir.library.oregonstate.edu/concern/datasets/2n49t8455>.

AUTHOR CONTRIBUTIONS

JM conceived of the field study. AH designed and collected all the data on the measured traits and contributed to significant portions of the writing through her master's thesis. AH and JM performed the analysis in SAS for the ANOVA, correlations, and heritabilities. LW condensed and edited the manuscript of AH, contributed significantly to the writing, performed the GWAS, rrBLUP, ANOVA, and QQ plots of residuals in R, and created LD heat maps and PCA in TASSEL. JM and LW performed genomic prediction analyses, and edited and contributed to all parts of the writing and analysis. All authors contributed to the article and approved the submitted version.

FUNDING

Funding for AH's graduate research assistantship, materials, and land use was provided by the Baggett-Frazier Vegetable Breeding and Genetics Endowment. Funding for development and genotyping of the Snap Bean Diversity Panel was provided by the Bean Coordinated Agricultural Project (USDA-NIFA-2009-85606-05964).

ACKNOWLEDGMENTS

This project was a contribution to the United States Department of Agriculture-National Institute of Food and Agriculture (USDA-NIFA) W3150 multistate project. We thank Shinji Kawai and Joel Davis for technical and logistical support for this project. The assistance of Kara Young, Zach Kraemer, and Ty Seely is gratefully acknowledged. We also thank Jennifer Kling for her assistance with the data analysis process and Samira Mafi-Moghaddam for training in GWAS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.697615/full#supplementary-material>

REFERENCES

- Abawi, G., Crosier, D., and Cobb, A. (1985). *Root rot of snap beans in New York. New York's Food and Life Sciences Bulletin*. Geneva, NY: New York State Agricultural Experiment Station.
- Abawi, G. S. (1990). *Root rots of beans in Latin America and Africa: Diagnosis, research methodologies, and management strategies*. Cali, Colombia: CIAT.
- Arenas, S., Cortés, A. J., Mastretta-Yanes, A., and Jaramillo-Correa, J. P. (2021). Evaluating the accuracy of genomic prediction for the management and conservation of relictual natural tree populations. *Tree Genet. Genom.* 17, 1–19.
- Assefa, T., Mahama, A. A., Brown, A. V., Cannon, E. K., Rubyogo, J. C., Rao, I. M., et al. (2019). A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*Phaseolus vulgaris* L.). *Mol. Breed.* 39:20. doi: 10.1007/s11032-018-0920-0
- Azzam, H. A. (1956). *Inheritance of resistance to Fusarium root rot in Phaseolus vulgaris L. and Phaseolus coccineus L.* Ph.D. dissertation. Corvallis, OR: Oregon State University.
- Baggett, J. R., Frazier, W. A., and Vaughan, E. K. (1965). Tests of *Phaseolus* species for resistance to *Fusarium* root rot. *Plant Dis. Rep.* 49:630.
- Bernardo, R. (2020). Reinventing quantitative genetics for plant breeding: Something old, something new, something borrowed, something BLUE. *Heredity* 125, 375–385.
- BIC. (2021a). *Bean Improvement Cooperative Research Techniques: Bean Root Rots*. Available Online at: http://www.bic.uprm.edu/?page_id=79 (verified 26-May-2021)
- BIC. (2021b). *Bean Improvement Cooperative: SCAR markers linked with disease resistance traits in common bean (Phaseolus vulgaris)* Updated: December 2010. Available Online at: http://arsftfbean.uprm.edu/bic/wp-content/uploads/2018/04/SCAR_Markers_2010.pdf (verified 15-Jun-21)
- Blair, M. W., Cortés, A. J., Penmetas, R. V., Farmer, A., Carrasquilla-Garcia, N., and Cook, D. R. (2013). A high-throughput SNP marker system for parental polymorphism screening, and diversity analysis in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 126, 535–548.
- Blair, M. W., Soler, A., and Cortes, A. J. (2012). Diversification and population structure in common beans (*Phaseolus vulgaris* L.). *PLoS One* 7:e49488. doi: 10.1371/journal.pone.0049488
- Burke, D. W., and Miller, D. E. (1983). Control of *Fusarium* root rot with resistant beans and cultural management. *Plant Dis.* 67, 1312–1317.
- Cao, X., Zhai, X., Xu, E., Zhao, Z., and Fan, G. (2020). Genome-wide identification of candidate genes related to disease resistance and high biomass in tetraploid *Paulownia*. *Acta Physiol. Plant.* 42:171. doi: 10.1007/s11738-020-03160-7
- Cavanagh, C., Morell, M., Mackay, I., and Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* 11, 215–221.
- Chowdhury, M. A., Yu, K., and Park, S. J. (2002). Molecular mapping of root rot resistance in common beans. *Annu. Rep. Bean Improv. Coop.* 45, 96–97.
- Cirak, M., and Myers, J. R. (2021). Cosmetic stay-green trait in snap bean and the event cascade that reduces seed germination and emergence. *J. Amer. Soc. Hort. Sci.* 2021, 1–11. doi: 10.21273/JASHS05038-20
- Cortés, A. J., Chavarro, M. C., and Blair, M. W. (2011). SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theor. Appl. Genet.* 123:827.
- Cortés, A. J., Skeen, P., Blair, M. W., and Chacón-Sánchez, M. I. (2018). Does the genomic landscape of species divergence in *Phaseolus* beans coerce parallel signatures of adaptation and domestication? *Front. Plant Sci.* 9:1816. doi: 10.3389/fpls.2018.01816
- Diaz, L. M., Arredondo, V., Ariza-Suarez, D., Aparicio, J., Buendia, H. F., Cajiao, C., et al. (2021). Genetic analyses and genomic predictions of root rot resistance in common bean across trials and populations. *Front. Plant Sci.* 12:629221. doi: 10.3389/fpls.2021.629221
- Diaz, S., Ariza-Suarez, D., Ramdeen, R., Aparicio, J., Arunachalam, N., Hernandez, C., et al. (2021). Genetic architecture and genomic prediction of cooking time in common bean (*Phaseolus vulgaris* L.). *Front. Plant Sci.* 11:622213. doi: 10.3389/fpls.2020.622213
- Dramadri, I. O., Amongi, W., Kelly, J. D., and Mukankusi, C. M. (2020). Genome-wide association analysis of resistance to *Pythium ultimum* in common bean (*Phaseolus vulgaris*). *Plant Breed.* 139, 1168–1180. doi: 10.1111/pbr.12855
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024
- Gao, X., Becker, L. C., Becker, D. M., Starmer, J. D., and Province, M. A. (2010). Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet. Epidemiol.* 34, 100–105.
- Goh, L., and Yap, V. B. (2009). Effects of normalization on quantitative traits in association test. *BMC Bioinform.* 10:415. doi: 10.1186/1471-2105-10-415
- Hagerty, C. H., Cuesta-Marcos, A., Cregan, P. B., Song, Q., McClean, P., Noffsinger, S., et al. (2015). Mapping *Fusarium solani* and *Aphanomyces euteiches* root rot resistance and root architecture quantitative trait loci in common bean. *Crop Sci.* 55, 1969–1977.
- Hallauer, A. R., Filho, J. B. M., and Carena, M. J. (2010). *Hereditary Variance: Mating Designs, In: Quantitative Genetics in Maize Breeding, Handbook of Plant Breeding*. New York: Springer, 81–167.
- Heslot, N., Yang, H.-P., Sorrells, M. E., and Jannink, J.-L. (2012). Genomic selection in plant breeding: A comparison of models. *Crop Sci.* 52, 146–160.
- Holland, J. B., Nyquist, W. E., and Cervantes-Martínez, C. T. (2003). Estimating and interpreting heritability for plant breeding: an update. *Plant Breed. Rev.* 22, 9–112.
- Jiang, S.-C., Mei, C., Liang, S., Yu, Y.-T., Lu, K., Wu, Z., et al. (2015). Crucial roles of the pentatricopeptide repeat protein SOAR1 in *Arabidopsis* response to drought, salt and cold stresses. *Plant Mol. Biol.* 88, 369–385. doi: 10.1007/s11103-015-0327-9
- Kamfwa, K., Mwala, M., Okori, P., Gibson, P., and Mukankusi, C. (2013). Identification of QTL for *Fusarium* root rot resistance in common bean. *J. Crop. Improv.* 27, 406–418. doi: 10.1080/15427528.2013.786775
- Keller, B., Ariza-Suarez, D., De La Hoz, J., Aparicio, J. S., Portilla-Benavides, A. E., Buendia, H. F., et al. (2020). Genomic prediction of agronomic traits in common bean (*Phaseolus vulgaris* L.) under environmental stress. *Front. Plant Sci.* 11:1001. doi: 10.3389/fpls.2020.01001
- Kleintop, A. E., Myers, J. R., Echeverria, D., Thompson, H. J., and Brick, M. A. (2016). Total phenolic content and associated phenotypic traits in a diverse collection of snap bean cultivars. *J. Amer. Soc. Hort. Sci.* 141, 3–11.
- Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12:e1005767. doi: 10.1371/journal.pgen.1005767
- Liu, X., Wang, H., Wang, H., Guo, Z., Xu, X., Liu, J., et al. (2018). Factors affecting genomic selection revealed by empirical evidence in maize. *Crop J.* 6, 341–352.
- López-Hernández, F., and Cortés, A. J. (2019). Last-generation genome-environment associations reveal the genetic basis of heat tolerance in common bean (*Phaseolus vulgaris* L.). *Front. Genet.* 10:22. doi: 10.3389/fpls.2018.00022
- Lynch, J., and Brown, K. (2013). *Common bean shovelomics*. Available Online at: <http://plantscience.psu.edu/research/labs/roots/methods/field/shovelomics/intensive-bean-crown-phenotyping> (verified 07-Apr-2021)
- Lynch, J. P., and Brown, K. M. (2001). Topsoil foraging - an architectural adaptation of plants to low phosphorus availability. *Plant Soil* 237, 225–237.
- Mamidi, S., Lee, R. K., Goos, J. R., and McClean, P. E. (2014). Genome-wide association studies identifies seven major regions responsible for iron deficiency chlorosis in soybean (*Glycine max*). *PLoS One* 9:e107469. doi: 10.1371/journal.pone.0107469
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorf, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753.
- Mukankusi, C., Derera, J., Melis, R., Gibson, P., and Buruchara, R. (2011). Genetic analysis of resistance to *Fusarium* root rot in common bean. *Euphytica* 182, 11–23. doi: 10.1007/s10681-011-0413-2
- Myers, J. R., Wallace, L. T., Mafi Moghaddam, S., Kleintop, A. E., Echeverria, D., Thompson, H. J., et al. (2019). Improving the health benefits of snap bean: genome-wide association studies of total phenolic content. *Nutrients* 11:2509. doi: 10.3390/nu11102509
- Nakedde, T., Ibarra-Perez, F. J., Mukankusi, C., Wainess, J. G., and Kelly, J. D. (2016). Mapping of QTL associated with *Fusarium* root rot resistance and root architecture traits in black beans. *Euphytica* 212, 51–63. doi: 10.1007/s10681-016-1755-6

- Navarro, F., Sass, M. E., and Nienhuis, J. (2008). Identification and confirmation of quantitative trait loci for root rot resistance in snap bean. *Crop Sci.* 48, 962–972. doi: 10.2135/cropsci2007.02.0113
- Oladzad, A., Zitnick-Anderson, K., Jain, S., Simons, K., Osorno, J. M., McClean, P. E., et al. (2019). Genotypes and genomic regions associated with *Rhizoctonia solani* resistance in common bean. *Front. Plant Sci.* 10:956. doi: 10.3389/fpls.2019.00956
- Paulino, J. F. D. C., Almeida, C. P. D., Bueno, C. J., Song, Q., Fritsche-Neto, R., Carbonell, S. A. M., et al. (2021). Genome-wide association study reveals genomic regions associated with fusarium wilt resistance in common bean. *Genes* 12:765. doi: 10.3390/genes12050765
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovič, M., et al. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *J. Evol. Biol.* 30, 1450–1477.
- Ray, H., Douches, D. S., and Hammerschmidt, R. (1998). Transformation of potato with cucumber peroxidase: Expression and disease response. *Physiol. Mol. Plant Path.* 53, 93–103.
- Resende, M. F. R., Muñoz, P., Resende, M. D. V., Garrick, D. J., Fernando, R. L., Davis, J. M., et al. (2012). Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). *Genetics* 190, 1503–1510.
- Riedelsheimer, C., Technow, F., and Melchinger, A. E. (2012). Comparison of whole-genome prediction models for traits with contrasting genetic architecture in a diversity panel of maize inbred lines. *BMC Genom.* 13:452. doi: 10.1186/1471-2164-13-452
- Román-Avilés, B., and Kelly, J. D. (2005). Identification of quantitative trait loci conditioning resistance to *Fusarium* root rot in common bean. *Crop Sci.* 45, 1881–1890.
- Scheet, P., and Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* 78, 629–644. doi: 10.1086/502802
- Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nat. Genet.* 46, 707–713. doi: 10.1038/ng.3008
- Schneider, K. A., Grafton, K. F., and Kelly, J. D. (2001). QTL Analysis of resistance to *Fusarium* root rot in bean. *Crop Sci.* 41:535. doi: 10.2135/cropsci2001.412535x
- Shenkut, A. A., and Brick, M. A. (2003). Traits associated with dry edible bean (*Phaseolus vulgaris*). *Euphytica* 133, 339–347. doi: 10.1023/A:1025774110004
- Shi, A., Gepts, P., Song, Q., Xiong, H., Michaels, T. E., and Chen, S. (2021). Genome-wide association study and genomic prediction for soybean cyst nematode resistance in USDA common bean (*Phaseolus vulgaris*) core collection. *Front. Plant Sci.* 12:624156. doi: 10.3389/fpls.2021.624156
- Silbernagel, M. J. (1987). *Fusarium* root rot-resistant snap bean breeding line FR-266. *HortScience* 22, 1337–1338.
- Snapp, S., Kirk, W., Román-Avilés, B., and Kelly, J. (2003). Root traits play a role in integrated management of *Fusarium* root rot in snap beans. *HortScience* 38, 187–191.
- Soltani, A., MafiMoghaddam, S., Oladzad-Abbasabadi, A., Walter, K., Kearns, P. J., Vasquez-Guzman, J., et al. (2018). Genetic analysis of flooding tolerance in an Andean diversity panel of dry bean (*Phaseolus vulgaris* L.). *Front. Plant Sci.* 9:767. doi: 10.3389/fpls.2018.00767
- Song, Q., Jia, G., Hyten, D. L., Jenkins, J., Hwang, E.-Y., Schroeder, S. G., et al. (2015). SNP assay development for linkage map construction, anchoring whole-genome sequence, and other genetic and genomic applications in common bean. *G3 (Bethesda)* 5, 2285–2290. doi: 10.1534/g3.115.020594
- Spindel, J. E., Begum, H., Akdemir, D., Collard, B., Redoña, E., Jannink, J. L., et al. (2016). Genome-wide prediction models that incorporate *de novo* GWAS are a powerful new tool for tropical rice improvement. *Heredity* 116, 395–408. doi: 10.1038/hdy.2015.113
- Thistlethwaite, F. R., Gamal El-Dien, O., Ratcliffe, B., Klápštil, J., Porth, I., Chen, C., et al. (2020). Linkage disequilibrium vs. pedigree: genomic selection prediction accuracy in conifer species. *PLoS One* 15:e0232201. doi: 10.1371/journal.pone.0232201
- Wallace, L. T., Arkwazee, H., Vining, K., and Myers, J. R. (2018). Genetic diversity within snap beans and their relation to dry beans. *Genes* 9:587. doi: 10.3390/genes9120587
- Wang, W., Jacobs, J. L., Chilvers, M. I., Mukankusi, C. M., Kelly, J. D., and Cichy, K. A. (2018). QTL analysis of *Fusarium* root rot resistance in an Andean × Middle American common bean RIL population. *Crop Sci.* 58, 1166–1180.
- Wen, L., Chang, H. X., Brown, P. J., Domier, L. L., and Hartman, G. L. (2019). Genome-wide association and genomic prediction identifies soybean cyst nematode resistance in common bean including a syntenic region to soybean *Rhg1* locus. *Hort. Res.* 6, 1–12. doi: 10.1038/s41438-018-0085-3
- Xi, W., ShiWen, W., Dan, C., DaoQian, C., Meijuan, Z., XiPing, D., et al. (2015). Influence of overexpression of 2-alkenal reductase gene on drought resistance in tobacco. *Acta Bot. Boreali-Occidentalia Sinica* 35, 1166–1172.
- Zitnick-Anderson, K., Oladzadabbasabadi, A., Jain, S., Modderman, C., Osorno, J. M., McClean, P. E., et al. (2020). Sources of resistance to *Fusarium solani* and associated genomic regions in common bean diversity panels. *Front. Genet.* 11:475. doi: 10.3389/fgene.2020.00475

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 Huster, Wallace and Myers. 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 and Validation of Genomic Regions Associated With Charcoal Rot Resistance in Tropical Maize by Genome-Wide Association and Linkage Mapping

Zerka Rashid¹, Harleen Kaur², Veerendra Babu¹, Pradeep Kumar Singh^{1†}, Sharanappa I. Harlapur³ and Sudha K. Nair^{1*}

OPEN ACCESS

Edited by:

Anna Maria Mastrangelo,
Council for Agricultural and
Economics Research (CREA), Italy

Reviewed by:

Sandra E. Branham,
Clemson University, United States
Ivan Schuster,
Yuan Longping High-tech Agriculture
Co., Ltd., China

*Correspondence:

Sudha K. Nair
Sudha.nair@cgiar.org

†Present address:

Pradeep Kumar Singh,
Corteva AgriScience Seeds India Pvt
Ltd. Madhapur, Hyderabad, India

Specialty section:

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

Received: 17 June 2021

Accepted: 30 August 2021

Published: 08 October 2021

Citation:

Rashid Z, Kaur H, Babu V, Singh PK,
Harlapur SI and Nair SK (2021)
Identification and Validation of
Genomic Regions Associated With
Charcoal Rot Resistance in Tropical
Maize by Genome-Wide Association
and Linkage Mapping.
Front. Plant Sci. 12:726767.
doi: 10.3389/fpls.2021.726767

¹ International Maize and Wheat Improvement Center (CIMMYT), ICRISAT Campus, Hyderabad, India, ² Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, India, ³ Department of Plant Pathology, University of Agricultural Sciences, Dharwad, India

Charcoal rot is a post-flowering stalk rot (PFSR) disease of maize caused by the fungal pathogen, *Macrophomina phaseolina*. It is a serious concern for smallholder maize cultivation, due to significant yield loss and plant lodging at harvest, and this disease is expected to surge with climate change effects like drought and high soil temperature. For identification and validation of genomic variants associated with charcoal rot resistance, a genome-wide association study (GWAS) was conducted on CIMMYT Asia association mapping panel comprising 396 tropical-adapted lines, especially to Asian environments. The panel was phenotyped for disease severity across two locations with high disease prevalence in India. A subset of 296,497 high-quality SNPs filtered from genotyping by sequencing was correcting for population structure and kinship matrices for single locus mixed linear model (MLM) of GWAS analysis. A total of 19 SNPs were identified to be associated with charcoal rot resistance with P -value ranging from 5.88×10^{-06} to 4.80×10^{-05} . Haplotype regression analysis identified 21 significant haplotypes for the trait with Bonferroni corrected $P \leq 0.05$. For validating the associated variants and identifying novel QTLs, QTL mapping was conducted using two $F_2:3$ populations. Two QTLs with overlapping physical intervals, qMSR6 and qFMSR6 on chromosome 6, identified from two different mapping populations and contributed by two different resistant parents, were co-located with the SNPs and haplotypes identified at 103.51 Mb on chromosome 6. Similarly, several SNPs/haplotypes identified on chromosomes 3, 6 and 8 were also found to be physically co-located within QTL intervals detected in one of the two mapping populations. The study also noted that several SNPs/haplotypes for resistance to charcoal rot were located within physical intervals of previously reported QTLs for *Gibberella* stalk rot resistance, which opens up a new possibility for common disease resistance mechanisms for multiple stalk rots.

Keywords: GWAS—genome-wide association study, linkage (QTL) mapping, haplotype analysis, charcoal rot, maize

INTRODUCTION

Maize is cultivated on more than 180 million hectares (M ha) globally, contributing ~50% [1,117 million metric tons (MMTs)] to the global grain production (Prasanna, 2018). Asian countries have shown rapid progress in maize production and productivity and are the second largest maize producers in the world with 31% share in global maize production (Zaidi et al., 2018). China produced nearly 260.95 MMT of maize by cultivating the maize area of 41.30 M ha during 2019 (FAO., 2021). The second prime maize producing country among Asian countries is India with an estimated maize area of ~9.03 M ha in 2019 with the maize production of 27.72 M Mt at a productivity of 3.07 t/ha (FAO., 2021). In Asia, a large portion of maize (~70% of total volume) is used by the feed industry (Prasanna, 2018), and the maize demand is always increasing due to the rise in population and socio-economic growth (Shiferaw et al., 2011). Apart from feed, maize is increasingly used in industries especially in food processing industry for making additives and sweeteners (Prasanna, 2018).

Despite the substantial growth rates in terms of cultivated maize area, production, and productivity in the last few years, maize in the south and southeast Asia is largely (80%) grown as a rainfed crop that is prone to the vagaries of monsoon rains, in addition to a number of biotic and abiotic stresses in this region (Zaidi et al., 2018). Abiotic stresses like drought, heat, and waterlogging are the main stresses that have a high impact on yield loss. Compounded with these, diseases have a huge impact on grain yield, as observed in most of the countries in Asia. The most common and economically important diseases in the region are soil-borne diseases like post-flowering stalk rots (PFSR) and banded leaf and sheath blight (BLSB) and foliar diseases like Turicum leaf blight (TLB), downy mildews (DM), common rust, and polysora rust. Due to the impact of climate change effects, maize stalk rots and ear rots are reported to become more severe and widespread (Prasanna et al., 2021). Stalk rots in maize are caused by many fungi and bacteria, most of which occur commonly in the fields and behave opportunistically by infecting senescing, injured, and stressed plants (Jackson-Ziems et al., 2014). Stalk rots caused by fungi are *Fusarium* stalk rot (FSR), *Gibberella* stalk rot (GSR), late wilt, Anthracnose stalk rot (ASR), *Diplodia* stalk rot (DSR), and charcoal rot (CR).

Macrophomina phaseolina (Tassi) Goid., which causes charcoal rot of maize, is economically one of the most important pathogens that have a wide host range, affecting more than 500 species of plants. Microsclerotia of *M. phaseolina* survive in the soil, and the infected plant remains serve as a basic source of infection for the crops. They are ubiquitous under raised soil temperature and low moisture conditions, and in moisture-less soil, they can exist for more than 10 months (Khan, 2007). Charcoal rot symptoms are distinguished by the appearance of a large number of minute black sclerotia on vascular bundles and inside the rind of the stalk, resulting in grayish black stalk color. Symptoms of charcoal rot are observed after plant reproductive growth, when the fungus spreads into the lower internode of the stalk causing soft stalk, premature drying of stalk, and lodging of plants (Khokhar et al., 2014), and hence, the economic impact

of the disease is high. Disease severity is exacerbated by low soil moisture, and higher soil and air temperature (Smith and Wyllie, 1999), which are serious constraints faced under smallholder farming conditions in climate-vulnerable environments. It is distributed worldwide in the tropics and subtropics, as well as in the US northern, central, and southern regions (Wyllie, 1988). It is a serious biotic concern in Asian countries like China, India, Indonesia, Pakistan, Philippines, Thailand, and Vietnam (Sharma et al., 1993). Yield loss due to charcoal rot was estimated to be 25–32.2% in India (Kumar et al., 1996) and recorded as high as 63.5% in All India Coordinated Research Program trials (Maize AICRP., 2014). These losses can be avoided by the deployment of resistant cultivars, as chemical control to soil-borne diseases has been reported as largely ineffective, and it increases the cultivation cost of resource-constrained farmers, apart from having hazardous effects on the environment.

Resistance to CR is shown to be a polygenic trait, with additive and non-additive gene action, with significant environmental interaction (Singh and Kaiser, 1991; Krishna et al., 2013; Mir et al., 2018). Incorporating resistance to diseases like charcoal rot, which are quantitatively inherited and have significant environmental interaction, in the breeding schemes to enhance genetic gains over time, necessitates the use of all modern breeding tools and strategies. Molecular technologies are used to accelerate the breeding for disease resistance by the possibility to expand the size of breeding populations, thereby increasing selection intensity, without increasing phenotyping requirements. Genotypic information can be used to select germplasm at the early stages of selection, and the capability to increase this phenotypically untested layer will allow the total number of genotypes within a breeding program to be expanded (Cooper et al., 2014). Linkage mapping can be used to identify quantitative trait loci (QTLs), which, in turn, are the tools for selection of loci of interest in breeding crosses and hence act as a proxy to the actual trait. Among the different PFSR, QTL mapping studies have been reported for resistance to stalk rots like GSR and ASR in maize. Three moderate to major QTLs have been identified, and one among them has been fine mapped for resistance to GSR caused by *Fusarium graminearum* (Yang, 2010; Zhang et al., 2012; Ma et al., 2017). A major QTL for ASR (caused by *Colletotrichum graminicola*) was cloned and found to belong to a nucleotide-binding site-leucine-rich repeat (NBS-LRR) gene class on the long arm of chromosome 4 (Jung et al., 1994; Abad et al., 2006; Broglie et al., 2011). Molecular mapping studies for charcoal rot resistance in maize have not been reported yet, which may be attributed to several factors like limited availability of disease-resistant sources, complex nature of the disease, and possible co-infection with other stalk rot pathogens under natural conditions leading to low repeatability in trials. However, QTLs for resistance to charcoal rot caused by *M. phaseolina* have been reported in crops like sorghum (Mahmoud et al., 2018), soybean (da Silva et al., 2019), and sesame (Wang et al., 2017). Keeping in view the increasing incidences of charcoal rot of maize in South Asia and gap in knowledge on the genomic regions conferring resistance to the trait, we conducted this study to discover trait markers through genome-wide association mapping and haplotype analysis using CIMMYT Asia association

mapping (CAAM) panel. The genomic regions associated with charcoal rot resistance identified were validated using QTL mapping in two mapping populations, apart from identifying population-specific QTLs. Validated regions/markers will be further studied in breeding populations for possible deployment in the breeding pipelines.

MATERIALS AND METHODS

Plant Material

A set of 396 lines from the CAAM panel that were developed and adapted in Asian environments, involving inbred lines with tolerance to abiotic stresses like drought, high temperature, and excess moisture, besides quality protein maize (QPM) lines, and inbred lines derived from downy mildew-resistant populations in Asia, was used in genome-wide association study (GWAS). The CAAM panel included lines that are adapted to tropical, subtropical, lowland, mid-altitude, and highland environments and was classified into early maturing, intermediate maturing, and late maturing based on growing degree days (GDD). Most of the lines had yellow/orange kernel color, with very few lines had white kernel color (**Supplementary Table 1**).

Two biparental $F_{2:3}$ families were formed to perform linkage mapping analysis for the validation of GWAS results. The first population (MSR) derived from a cross between a charcoal rot-resistant female parent CML495 and a susceptible male parent CML474 comprised 190 $F_{2:3}$ families. CML495 is an elite lowland adapted, late inbred line with white kernel color. The second population (FMSR) derived from a cross between a resistant female parent WLS-F36-4-2-2-B-1-B*9 (now released as CML578) and a susceptible male parent CML474 comprised 257 F_3 families. The common susceptible parent CML474 is an Asia-lowland adapted early line used as the early generation tester for heterotic group A.

Phenotypic Evaluation Screening Sites

The CAAM panel was evaluated under artificial inoculation conditions for charcoal rot at two hot spot locations: Borlaug Institute for South Asia (BISA) farm, Ludhiana, Punjab, India (30°55' N, 75°54' E; 229 masl; 750–800 mm/year rainfall) during the wet season of 2013 and International Crop Research Institute for Semi-Arid Tropics (ICRISAT) farm, Hyderabad, Telangana, India (17.53° N; 78.27° E; 545 masl; 784 mm/year average rainfall) during the dry season of 2013 and 2014. For linkage mapping, $F_{2:3}$ families of two mapping populations, MSR and FMSR, were evaluated for charcoal rot at ICRISAT farm, Hyderabad, during the dry season of 2017 and 2018, respectively. All disease evaluation trials were planted in alpha lattice design with two replications of a single row. The row length was 2 m with a spacing of 0.20 m between plant to plant and 0.75 m between row to row. Standard agricultural practices were maintained throughout the cropping season.

Inoculum Preparation and Inoculation Technique

Toothpick method was followed for artificial inoculation of the trials (Lal and Singh, 1984). In this method, mass multiplication

of *M. phaseolina* for artificial inoculation was done on wooden toothpicks by the method proposed by Jardine and Leslie (1992), with slight modifications. For inoculum multiplication, wooden toothpicks were saturated in tap water for 12–15 h followed by air drying. Dried toothpicks (~250) were packed in 250 ml glass bottles with 50 ml distilled water and were autoclaved at 15 lbs and 121°C for 15 min. After sterilization, excess water was poured out of the glass bottles and potato dextrose broth (PDB) was added, followed by autoclaving at the same temperature and pressure regime. After cooling, freshly subcultured fungi were inoculated into the bottles under aseptic conditions and incubated at 25°C till the toothpicks were covered up with fungal growth (~15 days).

At the tassel emergence stage of the plants, colonized toothpicks were inserted into the stalks. This was attained by drilling a hole of 4–5 cm at 45° angle in the second internode (first elongated node) with an iron needle having a wooden handle, where the toothpicks were introduced into the hole.

Disease Scoring

Disease scores were taken after 45–50 days of inoculation by splitting the stalk of the inoculated plants. Longitudinally divided stalks were individually scored on disease severity on a 1–9 scale (Payak and Sharma, 1983), where a score of 1 = 25% infection of the inoculated node; 2 = 26–50% of infection in the inoculated node; 3 = 51–75% of infection in the inoculated node; 4 = 76–100% of infection in the inoculated node; 5 = lesser than 50% of infection in the adjacent node; 6 = more than 50% of infection in the adjacent node; 7 = infection in more than three nodes; 8 = infection in more than four nodes; and 9 = infection in five nodes or plant lodging due to disease. Disease scores 1–2 were rated as highly resistant (HR), 2.1–4 were rated as resistant (R), 4.1–6 were rated as moderately resistant (MR), and >6.1 were rated as susceptible (S). In each row, at least 10 plants were inoculated, and each inoculated plant was scored to obtain a mean disease score for the plot.

Phenotypic Data Analysis

Descriptive statistics like mean, skewness, kurtosis, and genetic correlation were estimated using Meta-R (Alvarado et al., 2015). The CR disease data were skewed toward susceptibility in the CAAM panel. Best linear unbiased prediction (BLUPs) was obtained using the software Meta-6.0 across year data analysis, and the single year data were used for GWAS and QTL mapping analysis, respectively. The linear models are implemented in *lmer* from package *lme4* of R (R Core Team 2013) using REML to calculate BLUPs and estimate variance components. Broad-sense heritability of the combined analysis across years was estimated as $H^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2/e + \sigma_e^2/er)$, where σ_g^2 , σ_{ge}^2 , and σ_e^2 are the genotypic, genotype-by-year interaction, and error variance components, respectively, and e and r are the number of years and number of replicates within each year included in the corresponding analysis, respectively.

DNA Isolation and Genotyping of CAAM Panel

Genomic DNA of the maize lines in the association mapping panel was isolated from leaves of 3–4-week-old seedlings

(CIMMYT., 2001). Genotyping of the panel was performed at the Institute for Genomic Diversity, Cornell University, Ithaca, NY, USA, for single-nucleotide polymorphism (SNP) using genotyping by sequencing method (GBS). Genomic DNA was digested with the restriction enzyme *ApeKI*. The GBS libraries were constructed in 96-plex and sequenced in Illumina HiSeq 2000 (Elshire et al., 2011), and SNP calling was performed using TASSEL GBS pipeline (Glaubitz et al., 2014), where the GBS 2.7 sequences were used to anchor reads to the Maize B73 RefGen_v2 reference genome (www.maizgedb.org). Imputation was performed using FILLIN method in TASSEL 5.0, using GBS 2.7 haplotype files from Panzea (www.panzea.org) made from 8,000-site windows, as described in Swarts et al. (2014). The partially imputed GBS SNP data that had 955,690 genotypic data points (SNPs) across all the chromosomes were based on an algorithm that explores the closest neighbor in a small SNP window across the whole genome, permitting 5% mismatch (Romay et al., 2013). GWAS was conducted using 296,497 SNPs that were generated with the filtration criteria of call rate ≥ 0.7 and minor allele frequency (MAF) ≥ 0.05 .

GWAS and Haplotype Regression

Methods studied for GWAS analysis were naïve model, where genotypic data were used without correction (*G*-test); general linear model (GLM), where genotypic data were corrected for structure (*Q*) using 10 principal components (*G* + *Q*-test); and single locus mixed linear model (MLM), where genotypic data were corrected for both structure and kinship (*K*) (*G* + *Q* + *K*) to avoid spurious associations. Additive models were used for *G*-test and GLM, and mixed model single locus (EMMAX) (Kang et al., 2010) was used for MLM for association studies in SVS version 8.6.0 (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com). The mixed association mapping model used was $Y = \text{SNP} \cdot \beta + \text{PC} \cdot \alpha + K \cdot \mu + \epsilon$, where *Y* = response of the dependent variable (MSR Score), SNP = SNP marker (fixed effects), PC = principal component coordinate from the PCA (fixed effects), *K* = kinship matrix (random effects), α = vector of PC, β and μ = vectors of SNP and *K*, respectively, and ϵ = the error. A kinship matrix was estimated from identity-by-state distances matrix as executed in SVS version 8.6.0, where IBS distance = (no. of markers IBS2) + 0.5 × (no. of markers IBS1) no. of non-missing markers, where IBS1 and IBS2 are the states in which the two inbred lines share one or two alleles at a marker (Bishop and Williamson, 1990). Linkage disequilibrium (LD) was estimated on adjacent pairwise r^2 -values between adjacent SNPs among the SNPs from the GBS data and physical distances between those SNPs as described in Rashid et al. (2020). Manhattan and quantile-quantile plots were created using the association results. *P*-value threshold was estimated by using genome-wide LD between SNPs and the effective number of independent markers. Markers that were in approximate linkage equilibrium with each other were determined based on SNP pruning with LD r^2 threshold of 0.1 to select a subset of markers representing linkage blocks, and the suggestive-value threshold to control the genome-wide error rate was 5.16×10^{-5} (Mao et al., 2015; Cui et al., 2016). SNPs with $P \leq 0.01$ in GWAS of CAAM panel were selected for haplotype detection and trait regression. Expectation maximization (EM)

algorithm (Excoffier and Slatkin, 1995) with 50 EM iterations, EM convergence tolerance of 0.0001, and a frequency threshold of 0.01 were used to estimate haplotype frequency as applied in SVS version 8.6.0. Block defining algorithm (Gabriel et al., 2002) was used to identify haplotype blocks to minimize historical recombination. Regression analysis was carried out with the haplotype blocks identified on the MSR BLUP values based on stepwise regression with forward elimination.

Linkage Map Construction and Quantitative Trait Loci Mapping

Genomic DNA of the F_{2:3} lines of mapping population was extracted from the 3–4 weeks old seedlings. Markers were selected across the genome from the Illumina Goldengate assay for the QTL mapping study, apart from a few GWAS-identified SNPs. The lines were genotyped with KASP assays developed from random and GWAS-identified SNP markers at LGC Genomics, London. Based on parental line polymorphism, MSR mapping population was genotyped with a set of 125 markers, and the second population, FMSR, was genotyped with a set of 166 SNPs. Linkage map was constructed using QTL IciMapping version 3.4 software using the twin criterion of more than 3.0 LOD and a maximum distance of 40 cm between two loci. The QTLs were identified for BLUPs of the disease score using inclusive composite interval mapping (ICIM) as implemented in QTL IciMapping version 3.4. The walking step in QTL scanning was 1 cm, and a likelihood odds (LOD) threshold of 3.138 and 3.460 was used to declare QTL in MSR and FMSR populations, respectively, which was based upon 1,000 times permutations analysis. QTL statistics were also reported for those in which the LOD score exceeded 2.5. The sign of the additive effect of each QTL was used to identify the origin of the favorable allele.

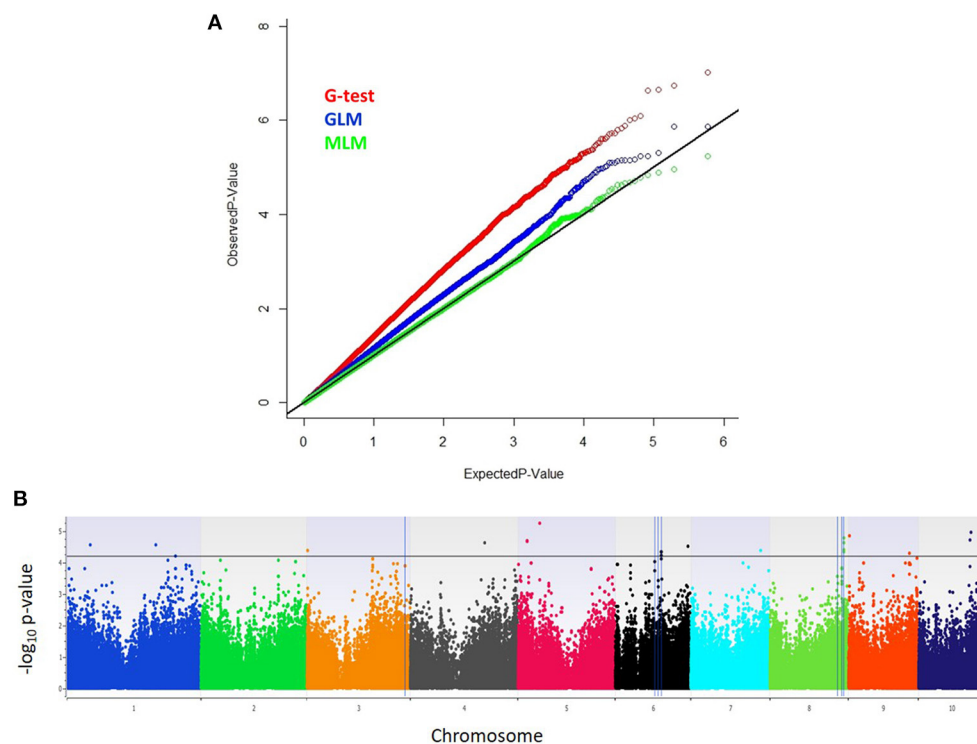
RESULTS

Phenotypic Evaluation for Charcoal Rot Resistance

The CAAM panel consisting of 396 inbred lines was screened for charcoal rot resistance across three locations/years in India. The panel showed elevated disease severity, with a maximum score of 9.00 on a scale of 1.00–9.00 during all 3 years at two locations. Minimum disease scores of 2.10, 2.00, and 3.77 were observed at BISA, Ludhiana and Hyderabad, during years 1 and 2, respectively. The average disease score across locations was 7.21, which was skewed toward susceptibility. Broad-sense heritability (h^2) was moderate to high (0.54–0.67) across single location with significant genotypic variance ($P \leq 0.001$). QTL mapping population, MSR, evaluated at Hyderabad showed a trial mean of 5.62, with minimum and maximum disease scores of 3.76 and 7.79, respectively. The second mapping population, FMSR, showed an average trial mean of 6.21, with minimum and maximum scores of 3.42 and 8.90, respectively. Heritability estimates of MSR and FMSR trials were high, with 0.65 in MSR and 0.71 in FMSR population (Table 1). The response of both mapping populations showed continuous distribution for CR disease severity ranging from disease resistant or tolerant to susceptible reaction (Supplementary Figure 1). BLUPs were

TABLE 1 | Summary statistics of CIMMYT Asia association mapping panel evaluated at three environments and two $F_{2:3}$ linkage mapping populations evaluated during the dry season of 2017 and 2018.

| Location/year | Mean | Min | Max | Phenotypic variance | Error variance | Genotypic variance | G × E variance | Heritability |
|---------------|------|------|------|---------------------|----------------|--------------------|----------------|--------------|
| ICRISAT-13 | 7.71 | 2.00 | 9.00 | 2.57 | 1.67 | 1.73** | – | 0.67 |
| ICRISAT-14 | 7.17 | 3.77 | 9.00 | 1.46 | 1.35 | 0.79** | – | 0.54 |
| BISA-13 | 6.71 | 2.10 | 9.00 | 1.23 | 1.01 | 0.73** | – | 0.59 |
| Across | 7.25 | 4.51 | 9.00 | 0.73 | 1.89 | 0.41** | 0.00016** | 0.57 |
| MSR-MP | 5.62 | 3.76 | 7.79 | 0.59 | 0.42 | 0.38** | – | 0.65 |
| FMSR-MP | 6.21 | 3.42 | 8.90 | 0.84 | 0.49 | 0.60** | – | 0.71 |

** $P < 0.001$.**FIGURE 1** | (A) Inflation depicted by Q–Q plots of observed vs. expected $-\log_{10} (P\text{-values})$ plots for charcoal rot using the naïve association model (G-test), GLM (G + Q), and MLM (G + Q + K); G = genotype (fixed), Q = 10 principal components (fixed), K = kinship matrix (random) for CAAM panel. (B) Highly significant SNPs identified from MLM model using Manhattan plot, plotted with the individual SNPs on the X-axis and $-\log_{10} P$ -value of each SNP on the Y-axis. The horizontal line showed the cutoff P -value, and the vertical line represents the identified QTLs and haplotype blocks in these regions for charcoal resistance.

estimated to further conduct GWAS for charcoal rot resistance in association mapping panel and linkage mapping analysis.

GWAS for Resistance to Charcoal Rot

From high density imputed 955K GBS genotypic data, a subset of 296,497 SNPs fulfilling the criteria of call rate ≥ 0.7 and MAF ≥ 0.05 was used for conducting GWAS analysis. The quantile–quantile (QQ) plot with observed against expected $-\log_{10} P$ -value revealed that highest genomic inflation was observed in Naïve or G-test association model, followed by general linear model (GLM) or G+Q model, where genomic inflation was controlled with population structure using first 10 principal components

(PCs). However, mixed linear model (MLM) or G + Q+K model corrected for both population structure (Q) and kinship (K) sighted minimum genomic inflation as noticed in the QQ plots (Figure 1). Therefore, highly significant associations for charcoal rot resistance in the CAAM panel were determined based on MLM analysis. The narrow-sense heritability for charcoal rot resistance due to the associated SNPs was found to be 0.53. The total number of SNPs identified to be linked with charcoal rot resistance was 19 with P -value ranging from 5.88×10^{-6} to 4.80×10^{-5} (Figure 1). The most significant association detected for resistance to charcoal rot was with SNP S5_48504604 on chromosome 5, which showed the lowest P -value, followed by

SNP S10_117560618 on chromosome 10. Among the 19 SNPs detected, groups of SNPs located at close physical co-ordinates were found on chromosome 5 (S5_19528704, S5_19528705), chromosome 6 (S6_103513337 and S6_103513378), and chromosome 8 (S8_165726551, S8_165726553, S8_165726556, and S8_165726574) (Table 2). Based on the physical position of the significant SNPs with respect to B73 version 2 of the reference genome (http://ensembl.gramene.org/Zea_mays), the significant SNPs identified in GWAS were associated with 12 genes, several of which had functional domains involved in resistance to biotic stresses.

Haplotype Detection and Regression Analysis

Two hundred and eighty-nine SNPs (with $P \leq 10^{-3}$) that were identified in GWAS analysis were used to construct 44 haplotype blocks across 10 chromosomes, which were used in haplotype regression (HTR) analysis on estimated BLUP values. HTR analysis identified 21 haplotypes with Bonferroni $P \leq 0.05$, which explained 3.22–6.48% of phenotypic variance. Haplotype blocks for charcoal rot resistance were identified on chromosomes 1, 2, 3, 5, 6, 8, and 9, formed with 2–8 SNPs (Table 3). Hap_8.1 on chromosome 8 formed by two SNPs, S8_151908973 and S8_151908983, showed the highest significance (Bonferroni P -value 7.73×10^{-05}), followed by the Hap_5.2 on chromosome 5 (P -value 5.11×10^{-06} and Bonferroni P -value 2.24×10^{-04}) (Table 3).

Linkage Mapping for Charcoal Resistance

Two biparental mapping populations phenotyped for charcoal rot at Hyderabad, India, were used for QTL mapping and validation of the genomic regions identified through GWAS and HTR analysis. By genotyping the MSR and FMSR populations with 125 and 166 markers, respectively, linkage maps were constructed. The average marker densities for MSR and FMSR populations were 7.09 and 5.59 cm, respectively, across the 10 chromosomes. Inclusive composite interval mapping in MSR mapping population identified two QTLs on chromosomes 6 and 8 (Figure 2), and two other QTLs were detected on chromosomes 3 and 4 at the lower default LOD threshold of 2.5. QTL qMSR8 on chromosome bin 8.06–07 between markers PZA01964_29 and PHM4757_14 had the largest effect, which explained 13.86% of the phenotypic variation. Resistant alleles were contributed by the resistant parent CML495 for all the QTLs identified in MSR population. In the FMSR mapping population, no QTLs were detected at the LOD threshold of 3.460, and two QTLs were identified on chromosomes 6 and 7 (Figure 2) at a lower default threshold of 2.5. QTL qFMSR6 on chromosome bin 6.03–04 between the markers PZA01029_1 and S6_103513510 showed the largest effect explaining 6.56% of the phenotypic variance (Table 4). For the two QTLs, resistant alleles were contributed by the resistant parent (WLS-F36-4-2-2-B-1-B*9). QTLs qMSR6 and qFMSR6, identified on chromosome 6, were found to be overlapping based on the physical coordinates, and this region was identified in both GWAS and HTR analysis also. QTLs detected in the two mapping populations predominantly showed

dominant effects; however, two QTLs detected in MSR mapping population showed additive effects for charcoal rot resistance.

DISCUSSION

Post-flowering stalk rots are complex diseases, due to collective infection with multiple soil-borne pathogens, intensified by abiotic stresses like drought and further compounded by secondary infections. Charcoal rot, caused by soil-borne pathogen *M. phaseolina*, is an important component of the PFSRs and its management methods include cultural practices, fungicide application, biological control, and resistant varieties. A comprehensive understanding of the host plant resistance is necessary to develop and deploy elite, stress-resistant varieties with little yield reduction in the presence of biotic stresses. As there are no reported studies on resistance to charcoal rot resistance in maize, we undertook this study to discover and validate genomic regions controlling this trait. A GWAS was conducted using a mapping panel that included tropical/subtropical inbred lines from CIMMYT breeding programs in Asia, Mexico, Kenya, Zimbabwe, and Colombia that are also acclimatized to the Asian tropics. The CAAM panel was previously used to study traits like resistance to sorghum downy mildew (Rashid et al., 2018), northern leaf corn blight (Rashid et al., 2020), and root traits under drought conditions (Zaidi et al., 2016) in Asian environments. Phenotypic evaluation of CAAM panel for charcoal rot at Hyderabad and Ludhiana, revealed that the panel was skewed toward susceptibility, possibly because both these locations had ideal environment for pathogen infection and spread, and the artificial inoculation using the toothpick method reduced the chances of escapes. The toothpick method has been widely used for artificial inoculation of stalk rots due to its simplicity and low cost (Tesso et al., 2009). In this study, we used linkage mapping apart from GWAS to study the genomic regions conferring resistance to charcoal rot. The high disease score mean in the AM panel compared with the mapping populations showed that the allele frequency of the resistant alleles might be lesser in the AM panel, whereas in the mapping populations such alleles contributed by the resistant parents were segregating in the populations, and hence higher allele frequency and lesser disease incidence.

Linkage mapping targets genetic recombination generated in artificially controlled crosses and offers huge advantages in terms of QTL detection power. However, it has the disadvantages of low mapping resolution, allele sampling, and speed. Unlike linkage mapping, GWAS makes use of the ancestral recombination events in a natural population to analyze marker-phenotype relations (Rafalski, 2002). It has the advantage of increased mapping resolution and speed but could have a lesser power of mapping (Korte and Farlow, 2013). Whereas, QTL mapping in biparental populations segregating for the relevant alleles at the associated/linked locus may be used in the validation of trait association (Rafalski, 2010), it also identifies novel QTLs not identified in GWAS, if the alleles are rare in the AM panel and/or the allelic phase differs across population structure groups (Famoso et al., 2011). To complement the GWAS analysis carried

TABLE 2 | Significantly associated single-nucleotide polymorphisms (SNPs) along with the predicted gene model and their function detected by genome-wide association studies in CIMMYT association mapping panel for charcoal rot resistance.

| Marker | Ch | P-Value | PVE% | Favorable allele | Predicted gene model | Gene name/best matching ortholog | Plants | Reported function | References |
|---------------|----|------------------------|------|------------------|----------------------|--|-----------------------------------|---|--|
| S5_48504604 | 5 | 5.88×10^{-06} | 5.63 | G | – | – | – | – | |
| S10_117560618 | 10 | 1.12×10^{-05} | 5.3 | A | GRMZM2G072513 | OSJNBa0088K19.7-Rice like protein | – | – | |
| S10_144684808 | 10 | 1.29×10^{-05} | 5.22 | G | GRMZM2G136895 | <i>Zea mays</i> Beta-D-xylosidase 4 | <i>Arabidopsis</i> , other plants | Cell wall modification, fruit development | Itai et al., 2003; Minic et al., 2004; Liao et al., 2012 |
| S9_1994787 | 9 | 1.44×10^{-05} | 5.17 | G | GRMZM2G500051 | – | – | – | |
| S8_165726556 | 8 | 1.68×10^{-05} | 5.09 | C | GRMZM2G414696 | – | – | – | |
| S10_115937334 | 10 | 1.97×10^{-05} | 5.01 | A | GRMZM2G050647 | Exocyst complex component SEC5 | <i>Arabidopsis</i> , other plants | Plant-pathogen interaction | Du et al., 2018 |
| S5_19528704 | 5 | 2.11×10^{-05} | 4.97 | G | GRMZM2G178767 | <i>Zea mays</i> Dof zinc-finger protein DOF5.7 | Plants | Abiotic stress, biotic stress | Sakamoto et al., 2004; Guo et al., 2009 |
| S5_19528705 | 5 | 2.19×10^{-05} | 4.95 | A | | | | | |
| S4_167190764 | 4 | 2.37×10^{-05} | 4.92 | A | GRMZM2G168337 | <i>Zea mays</i> Nicastrin | <i>Arabidopsis</i> , maize | Promotes maturation and proper trafficking of complex components and substrate recognition, biotic stress | Wang et al., 2012; Smolarkiewicz et al., 2014 |
| S8_165726551 | 8 | 2.38×10^{-05} | 4.91 | C | GRMZM2G414696 | – | – | – | – |
| S1_52605386 | 1 | 2.85×10^{-05} | 4.82 | T | – | – | – | – | – |
| S1_200489143 | 1 | 2.90×10^{-05} | 4.81 | T | GRMZM2G557453 | – | – | – | – |
| S6_163106367 | 6 | 3.07×10^{-05} | 4.78 | A | AC206312.3_FGT008 | – | – | – | – |
| S8_165726574 | 8 | 3.98×10^{-05} | 4.65 | A | GRMZM2G414696 | – | – | – | – |
| S3_2125663 | 3 | 4.31×10^{-05} | 4.61 | T | GRMZM2G170047 | <i>Zea mays</i> Cytochrome P450 71A26 | Maize, wheat, barley | Oxidation-reduction reaction, defense mechanism, secondary metabolite synthesis, Fusarium head blight | Morant et al., 2003; Irmisch et al., 2015; Gunupuru et al., 2018 |
| S7_156114994 | 7 | 4.34×10^{-05} | 4.61 | G | GRMZM2G465999 | <i>Zea mays</i> G-type lectin S-receptor-like serine/threonine-protein kinase B120 | Plants | Biotic and abiotic stress tolerance, plant defense | Lannoo and Van Damme, 2014 |
| S6_103513378 | 6 | 4.62×10^{-05} | 4.57 | C | GRMZM2G122172 | Aldehyde dehydrogenase family 2 member C4 | Plants | Abiotic and biotic stresses tolerance | Wen et al., 2012; Brocker et al., 2013 |
| S6_103513337 | 6 | 4.73×10^{-05} | 4.56 | A | | | | | |
| S8_165726553 | 8 | 4.80×10^{-05} | 4.55 | A | GRMZM2G414696 | – | – | – | |

Ch, chromosome, PVE, Phenotypic variance explained.

TABLE 3 | Significant haplotypes identified in the CAAM panel for resistance to charcoal rot using haplotype regression.

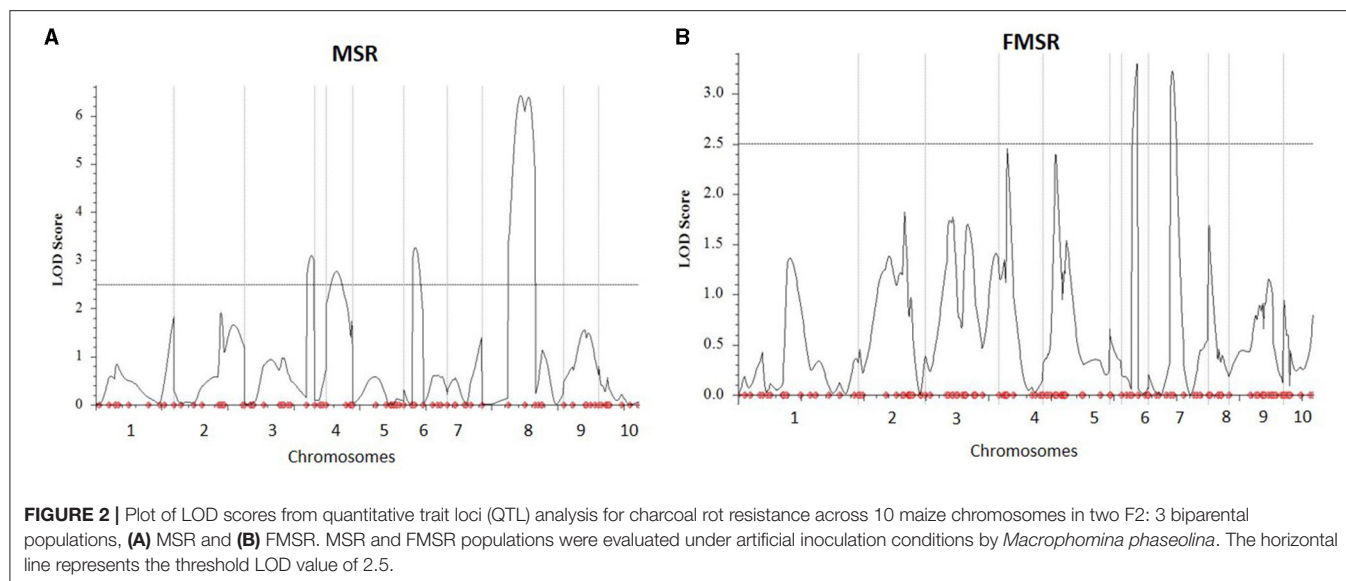
| Haplotype block | Ch | Markers | P-Value | PVE% | Bonferroni P-value | FDR | Favorable alleles |
|-----------------|----|--|-------------|-----------|--------------------|----------|-------------------|
| Hap_1.1 | 1 | S1_228148457, S1_228148501, S1_228148728, S1_228148775 | 0.000394055 | 4.4725128 | 0.017338434 | 0.001238 | TCGG |
| Hap_1.2 | | S1_245219474, S1_245219477, S1_245221215, S1_245221351, S1_245221353, S1_245221363 | 0.000434489 | 4.14338 | 0.019117509 | 0.001275 | CCCAAG |
| Hap_1.3 | | S1_259778254, S1_259778269 | 3.49E-05 | 4.6908192 | 0.001537657 | 0.000384 | CG |
| Hap_1.4 | | S1_268948971, S1_268948972, S1_268948974, S1_268948981 | 0.000289605 | 4.3917062 | 0.012742613 | 0.001274 | ATCC |
| Hap_1.5 | 2 | S1_290819006, S1_290819008 | 8.31E-06 | 5.8851488 | 0.000365777 | 0.000122 | GA |
| Hap_2.1 | | S2_212365693, S2_212365694 | 0.000543506 | 3.8516348 | 0.023914282 | 0.001407 | AG |
| Hap_3.1 | | S3_148298879, S3_148298896, S3_148298906, S3_148299049 | 0.000180157 | 4.3758201 | 0.007926917 | 0.000991 | GCCG |
| Hap_3.2 | | S3_168367332, S3_168367335, S3_168367337 | 0.00053713 | 3.7499107 | 0.02363372 | 0.001477 | CGG |
| Hap_3.3 | 3 | S3_202114642, S3_202114644 | 0.000666384 | 3.2229271 | 0.029320889 | 0.001466 | GC |
| Hap_3.4 | | S3_220734668, S3_220734677 | 0.000547048 | 3.3326993 | 0.024070091 | 0.001337 | CA |
| Hap_5.1 | | S5_19528704, S5_19528705, S5_19590454 | 4.59E-05 | 6.3666231 | 0.002019445 | 0.000404 | CCA |
| Hap_5.2 | | S5_68423958, S5_68423980 | 5.11E-06 | 5.8413714 | 0.000224844 | 0.000112 | CG |
| Hap_5.3 | 5 | S5_194559998, S5_194560001, S5_194560045, S5_194560047, S5_194560048 | 0.001015383 | 3.2534765 | 0.044676862 | 0.002127 | AATTA |
| Hap_6.1 | | S6_95934506, S6_95934536 | 0.000332882 | 3.5967653 | 0.014646826 | 0.001127 | GT |
| Hap_6.2 | | S6_103513337, S6_103513340, S6_103513378, S6_103513510 | 0.00017213 | 4.3611641 | 0.00757371 | 0.001082 | CGGG |
| Hap_8.1 | | S8_151908973, S8_151908983 | 1.76E-06 | 6.4890618 | 7.73E-05 | 7.73E-05 | GT |
| Hap_8.2 | 8 | S8_161523161, S8_161523199, S8_161523202, S8_161523204, S8_161523205, S8_161523207, S8_161523208, S8_161523210 | 0.000271353 | 4.1801244 | 0.011939511 | 0.001327 | GACTCTCT |
| Hap_9.1 | | S9_24597525, S9_24597528, S9_24597531, S9_24597534 | 4.96E-05 | 4.5259142 | 0.002180251 | 0.000363 | CTTG |
| Hap_9.2 | | S9_34173064, S9_34173069, S9_34173103 | 0.000311075 | 4.3479104 | 0.013687318 | 0.001141 | GGC |
| Hap_9.3 | | S9_137258399, S9_137258400, S9_137258402 | 0.00057319 | 3.4234421 | 0.025220351 | 0.001327 | GTG |
| Hap_9.4 | 9 | S9_137258446, S9_137258482 | 0.000291508 | 4.0121756 | 0.012826349 | 0.001166 | TA |

Ch, chromosome, PVE, Phenotypic variance explained, FDR, False discovery rate.

out in the Asia-adapted AM panel, two mapping populations, MSR and FMSR, with a common susceptible parent were used for linkage mapping to identify novel QTLs and for the validation of detected marker associations. A common susceptible early maturing parent, CML474, was used in both the mapping populations because it is highly susceptible to this disease and is being used as a tester for early maturity heterotic pool A. Markers spread across the genome were used for the QTL mapping study, along with some GWAS-identified SNP-based markers. There was no prior information on the status of the QTLs present in either of the parents. We conducted inclusive composite interval mapping to detect trait QTLs that were contributed by either of the parents. The QTLs identified that were co-located with the SNPs identified in GWAS were considered as validated in independent studies. The GWAS-SNPs that were

not co-located within QTL intervals were not considered as unvalidated, as they might just not be segregating in the parental combinations studied.

Two QTLs were detected with PVE ranging from 5.65 to 13.86% in one of the populations. Apart from these, three QTLs were detected at a lower threshold from the two populations. The results indicated that phenotypic variation for charcoal rot resistance in the two populations was explained largely by minor to moderate effect alleles. This is in accordance with QTL studies of a number of complex traits in maize. Out of the two QTLs identified in MSR populations, qMSR6 was found to overlap, based on physical co-ordinates, with qFMSR6, a minor QTL identified at a lower threshold of 2.5 in the FMSR population. This assumes immense significance as it is not very common to observe stable QTLs for complex traits contributed by unrelated



parental lines in different experiments. Among the genomic regions identified in GWAS, two SNPs on chromosome 6.03 (S6_103513337, S6_103513378) co-located with these QTLs were detected in linkage mapping. The trait-associated SNPs were located in the gene GRMZM2G122172, having the functional domain of aldehyde dehydrogenase (ALDH) family 2 member C4 (Carbon 4). Studies showed that ALDH upregulation is a common target of stress response pathway activation in plants, where ALDH responds to abiotic stresses leading to altered expression under exposure to stresses like drought (Bartels and Sunkar, 2005; Kotchoni et al., 2006). This function implies direct significance under charcoal rot infection, as the disease severity is directly related to drought stress. Also, studies have shown that ALDH gene from Chinese wild grapevine enhanced resistance to mildew pathogens and salt stress in *Arabidopsis* (Wen et al., 2012). Thus, this region can be considered as a region of interest for charcoal resistance and calls for further studies on dissection of the QTL and possible use in breeding populations. Another haplotype block, Hap_6.1 located at 95.93 Mb, was also identified within the QTL interval of qMSR6 identified in MSR mapping population.

On chromosome bin 8.06, four SNPs, S8_165726556, S8_165726551, S8_165726574, and S8_165726553, were identified, which co-located with the largest QTL identified in this study, qMSR8, located in the physical interval of 151.45 to 166.98 Mb on chromosome 8. Haplotype regression analysis identified two significant haplotypes (Hap_8.1 and Hap_8.2) for this trait within this QTL interval. In published studies on resistance to GSR, a major QTL *Rgsr8.1* was also fine mapped to 2.04 Mb interval between 164.69 and 166.72 Mb, with two candidate resistant genes, one of which was an auxin-responsive element and the other encoding a disease resistance protein (Chen et al., 2017). Also, co-incident with Hap_8.1, Ma et al. (2017) identified a QTL for resistance to GSR at physical position between 146.4 and 158.9 Mb. It is interesting to note that our

study identified and validated a genomic region contributing for resistance to charcoal rot, which also houses QTLs for resistance to another stalk rot pathogen, *Fusarium graminearum*, causing GSR. Apart from this, chromosomal bin 8.05–8.06 is known to harbor genes for resistance to multiple biotic stresses and is considered as one of the “complex, important and interesting” genomic regions in terms of maize disease resistance (Chung et al., 2010). Similar to this region on chromosome 8, trait-associated SNPs were identified on other chromosomes too that were located within previously mapped QTL intervals for GSR. Hap_3.4 was located within the QTL interval of minor QTL qMSR3 on chromosomal bin 3.09 at 220.73 Mb, where Ma et al. (2017) identified a QTL between 217.9 and 225.6 Mb for GSR resistance. In the same study, a major QTL *qRfg3*, at a physical position of 176.8–209.9 Mb, was detected across three field trials on chromosome bin 3.6/07 explaining 10.7–19.4% phenotypic variance for GSR resistance. The haplotype regression analysis for charcoal rot in this study identified Hap_3.3 on chromosome bin 3.07 at a physical position of 202.11 Mb, which fell within the QTL interval of QTL *qRfg3*. Ma et al. (2017) also identified a QTL on chromosome 5 between 49.9 and 152.0 Mb for GSR resistance, which also housed Hap_5.2 located at a physical co-ordinate of 68.42 Mb in this study. Further studies on gene characterization at these loci for both these diseases will be required to understand if common resistance mechanisms operate toward resistance to multiple stalk rot pathogens.

Several significant trait-associated SNPs identified in the GWAS were located within genes with functional domains related to biotic and abiotic stress tolerance, immune response, metabolism, plant development, and maturity (Table 2). Two SNPs, S5_19528704 and S5_19528705 identified for charcoal resistance and located in the same chromosomal bin as a QTL identified for GSR resistance on chromosome 5.02–5.04 (Pè et al., 1993) were located within the predicted gene GRMZM2G178767 that codes for a *Zea mays* Dof zinc-finger protein Dof5.7, which

TABLE 4 | Quantitative trait loci (QTL) detected on different chromosomes by inclusive composite interval mapping analysis for resistance to charcoal rot in two F_{2:3} biparental mapping populations.

| Population | QTL | Ch | Ch bin | Position, cm | Marker interval | Physical position (B73_V2) (LM) | Physical position (RM) | LOD | PVE (%) | Total PVE% | Additive effect | Dominant effect | D/A | Gene action |
|------------|--------------|----------|------------------|--------------|--------------------------------|---------------------------------|------------------------|-------------|--------------|------------|-----------------|-----------------|--------------|-------------|
| MSR | qMSR3 | 3 | 3.09 | 101 | PZA03391_1–ZA00316_10 | 219,859,920 | 223,513,639 | 3.1 | 5.72 | 23.54 | −0.1724 | 0.007 | −0.04 | A |
| | qMSR4 | 4 | 4.02–4.03 | 15 | PHM3963_33–PHM259_7 | 5,459,125 | 14,326,036 | 2.77 | 6.49 | | −0.143 | −0.1522 | 1.06 | D |
| | qMSR6 | 6 | 6.02–6.03 | 17 | PHM12904_7–S6_103513378 | 88,691,499 | 103,513,378 | 3.26 | 5.65 | | −0.1685 | 0.0012 | −0.01 | A |
| | qMSR8 | 8 | 8.06–8.07 | 59 | PZA01964_29–PHM4757_14 | 166,984,405 | 151,452,567 | 6.42 | 13.86 | | −0.2651 | −0.0721 | 0.27 | PD |
| FMSR | qFMSR6 | 6 | 6.03–6.04 | 26 | PZA01029_1–S6_103513510 | 114,031,392 | 103,513,510 | 3.298 | 6.56 | 10.48 | −0.2115 | −0.0393 | 0.185 | D |
| | qFMSR7 | 7 | 7.03 | 37 | PZA02643_1–PZA03166_1 | 128,365,318 | 137,632,654 | 3.223 | 6.51 | | −0.2068 | 0.0357 | −0.17 | D |

Ch, chromosome; LOD, likelihood odds; PVE, phenotypic variation explained; A, additive effect; D, dominant effect; PD, partially dominant.

The QTLs represented in bold show QTLs that had LOD scores above the LOD threshold based on 1,000 permutations, and the ones in normal fonts are the ones that were above LOD threshold of 2.5.

is implicated in abiotic and biotic stress tolerance in plants (Guo et al., 2009; Sakamoto et al., 2004). Zinc-finger domain is present in a well-known class of plant-resistant proteins, NBS-LRR, that are involved in effector-triggered immune response (Gupta et al., 2012). Zinc-finger-based WRKY transcription factor (TF) plays a broad and pivotal role in plant immune responses (Eulgem et al., 2007). Another significant SNP, S3_2125663, was located in the gene GRMZM2G170047 that potentially codes for cytochrome P450, which is known to boost disease resistance. Cytochrome P450s are membrane-bound enzymes that can accomplish oxidation–reduction reactions (Morant et al., 2003) and are involved in plant defense and secondary metabolite synthesis in classical xenobiotic detoxification pathway (Schuler and Werck-Reichhart, 2003). It was also reported to play a major role in resistance to *Fusarium* head blight disease caused by *Fusarium graminearum* in wheat (Walter et al., 2008; Walter and Doohan, 2011). Similarly, the gene GRMZM2G168337, which houses SNP S4_167190764, was implicated in the synthesis of Nicastrin, which was found to be upregulated in maize after inoculation with southern corn rust (Wang et al., 2012). A gene where a charcoal rot-associated SNP S7_156114994 was located is GRMZM2G465999, which is a type of lectin S-receptor-like serine/threonine-protein kinase. Plant kinases constitute a diverse protein superfamily, which is capable of recognizing and interacting with specific carbohydrate structures either from invading microorganisms or deformed plant cell wall structures, and plant lectin motifs are used constantly to combat against pathogens and predators during plant defense (Lannoo and Van Damme, 2014). Gene GRMZM2G050647 associated with SNP S10_115937334 codes for exocyst complex component SEC5, which plays a role in plant–pathogen interaction. Exocyst complex is a conserved multiprotein complex that has eight subunits that are used in pathogen defense against hemi-biotrophic pathogens like *Phytophthora infestans* and *Pseudomonas syringae*, and some exocyst subunits can act as a susceptibility factor for necrotrophic pathogens like *Botrytis cinerea*. (Du et al., 2018). In *Arabidopsis*, Exo70B mutants showed lesion-mimic cell death mediated by salicylic acid accumulation (Kulich et al., 2013).

CONCLUSION

The genetic architecture of charcoal rot resistance was dissected through association and linkage mapping. Nineteen SNPs were found to be highly significant for charcoal rot resistance in GWAS analysis, and haplotype regression identified 21 haplotypes, of which Hap_8.1 at 151.90 Mb on chromosome 8 was shown to have the most significant effect on the trait. Inclusive composite interval mapping in two F_{2:3} mapping populations detected QTLs on chromosomes 6 and 8 with PVE ranging from 5.65 to 13.86%. QTLs on chromosome bin 6.03, with a flanking marker at 103.51 Mb, were detected in both the linkage mapping populations, albeit at a lower threshold in one of the populations. SNPs/haplotypes in this QTL interval were identified in the GWAS and haplotype regression studies also. Similarly, the SNPs and haplotype detected on chromosome 8

were also validated in QTL mapping in one mapping population. These haplotypes on chromosomes 6 and 8 can be further analyzed in breeding populations for the possible deployment of trait markers for charcoal rot resistance. Several significant SNPs and haplotypes identified in this study were found to be located within published QTL intervals for GSR resistance. To our understanding, this study is the first report for mapping and validating genomic regions for charcoal rot resistance in maize.

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 at: <https://data.cimmyt.org/dataset.xhtml?persistentId=hdl:11529/10548606>.

REFERENCES

- Abad, L., Wolters, P., Stucker, D., and Davis, P. (2006). "Advances in anthracnose stalk rot resistance" in *The Fifth National IPM Symposium "Delivering on a Promise."* (St. Louis).
- Alvarado, G., López, M., Vargas, M., Pacheco, Á., Rodríguez, F., Burgueño, J., et al. (2015). *META-R (Multi Environment Trial Analysis With R for Windows) Version 4.1*. Available online at: <http://hdl.handle.net/11529/10201>.
- Bartels, D., and Sunkar, R. (2005). Drought and salt tolerance in plants. *Crit. Rev. Plant Sci.* 24, 23–58. doi: 10.1080/07352680590910410
- Bishop, D. T., and Williamson, J. A. (1990). The power of identity-by-state methods for linkage analysis. *Am. J. Hum. Genet.* 46, 254–265.
- Brocker, C., Vasilou, M., Carpenter, S., Carpenter, C., Zhang, Y., Wang, X., et al. (2013). Aldehyde dehydrogenase (ALDH) superfamily in plants: gene nomenclature and comparative genomics. *Planta* 237, 189–210. doi: 10.1007/s00425-012-1749-0
- Brogie, K. E., Butler, K. H., Butruille, M. G., Conceição, A. S., Frey, T. J., Hawk, J. A., et al. (2011). *Method for Identifying Maize Plants With RCG1 Gene Conferring Resistance to Colletotrichum Infection*. US patent 8,062,847.
- Chen, Q., Song, J., Du, W. P., Xu, L. Y., Jiang, Y., Zhang, J., et al. (2017). Identification, mapping, and molecular marker development for Rgsr8.1: a new quantitative trait locus conferring resistance to gibberella stalk rot in maize (*Zea mays* L.). *Front. Plant Sci.* 3:1355. doi: 10.3389/fpls.2017.01355
- Chung, C. L., Jamann, T., Longfellow, J., and Nelson, R. (2010). Characterization and fine mapping of a resistance locus for northern leaf blight in maize bin 8.06. *Theor. Appl. Genet.* 121, 205–227. doi: 10.1007/s00122-010-1303-z
- CIMMYT. (2001). *Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory Protocols*. Mexico: CIMMYT.
- Cooper, M., Messina, C. D., Podlich, D., Totir, L. R., Baumgarten, A., Hausmann, N. J., et al. (2014). Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop Pasture Sci.* 65, 311–336. doi: 10.1071/CP14007
- Cui, Z., Luo, J., Qi, C., Ruan, Y., Li, J., Zhang, A., et al. (2016). Genome-wide association study (GWAS) reveals the genetic architecture of four husk traits in maize. *BMC Genomics* 17:946. doi: 10.1186/s12864-016-3229-6
- da Silva, M. P., Pereira, A., Rupe, J. C., Bluhm, J. C., Gbur, E. E., Wood, L., et al. (2019). Effectiveness of a seed plate assay for evaluating charcoal rot resistance in soybean and the relationship to field performance. *Plant Dis.* 103, 1947–1953. doi: 10.1094/PDIS-10-18-1908-RE
- Du, Y., Overdijk, E. J. R., Berg, J. A., Govers, F., and Bouwmeester, K. (2018). Solanaceous exocyst subunits are involved in immunity to diverse plant pathogens. *J. Exp. Bot.* 69, 655–666. doi: 10.1093/jxb/erx442
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., and Buckler, E. S. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379. doi: 10.1371/journal.pone.0019379

AUTHOR CONTRIBUTIONS

SN designed the experiment. ZR, HK, VB, PS, and SH generated the phenotyping data. ZR and VB analyzed the data. SN and ZR wrote the manuscript. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support received from the CGIAR Research Program (CRP) on Maize.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.726767/full#supplementary-material>

- Eulgem, T., Imre, E., and Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Curr. Opin. Plant Biol.* 10:366–371. doi: 10.1016/j.pbi.2007.04.020
- Excoffier, L., and Slatkin, M. (1995). Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.* 12, 921–927.
- Famoso, A. N., Zhao, K., Clark, R. T., Tung, C. W., Wright, M. H., Bustamante, C., et al. (2011). Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLOS Genet.* 7:e1002221. doi: 10.1371/journal.pgen.1002221
- FAO. (2021). *FAOSTAT*. Rome: Food and Agriculture Organization of the United Nations. Available online at: <http://www.fao.org/faostat/en/#data>
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., et al. (2002). The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229. doi: 10.1126/science.1069424
- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS ONE* 9:e90346. doi: 10.1371/journal.pone.0090346
- Gunupuru, L. R., Arunachalam, C., Malla, K. B., Kahla, A., Perochon, A., Jia, J., et al. (2018). A wheat cytochrome P450 enhances both resistance to deoxynivalenol and grain yield. *PLoS ONE* 13:e0204992. doi: 10.1371/journal.pone.0204992
- Guo, Y. H., Yu, Y. P., Wang, D., Wu, C. A., Yang, G. D., Huang, J. G., et al. (2009). GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with GZIRD21A and GZIPR5. *New Phytol.* 183, 62–75. doi: 10.1111/j.1469-8137.2009.02838.x
- Gupta, S. K., Rai, A. K., Kanwar, S. S., and Sharma, T. R. (2012). Comparative analysis of zinc finger proteins involved in plant disease resistance. *PLoS ONE* 7:e42578. doi: 10.1371/journal.pone.0042578
- Irmisch, S., Zeltner, P., Handrick, V., Gershenzon, J., and Koller, T. G. (2015). The maize cytochrome P450 CYP79A61 produces phenylacetaldoxime and indole-3-acetaldoxime in heterologous systems and might contribute to plant defense and auxin formation. *BMC Plant Biol.* 15:128. doi: 10.1186/s12870-015-0526-1
- Itai, A., Ishihara, K., and Bewley, J. D. (2003). Characterization of expression, and cloning, of β -D-xylosidase and α -L-arabinofuranosidase in developing and ripening tomato (*Lycopersicon esculentum* mill.) fruit. *J. Exp. Bot.* 54, 2615–2622. doi: 10.1093/jxb/erg291
- Jackson-Ziems, T. A., Rees, J. M., and Harveson, R. M. (2014). *Common Stalk Rot Diseases of Corn*. Lincoln: DigitalCommons, University of Nebraska. Available online at: <https://digitalcommons.unl.edu/plantpathpapers/532>
- Jardine, D. J., and Leslie, J. F. (1992). Aggressiveness of *Gibberella fujikuroi* (*Fusarium moniliforme*) isolates to grain sorghum under greenhouse conditions. *Plant Dis.* 76, 897–900. doi: 10.1094/PD-76-0897
- Jung, M., Weldekidan, T., Schaff, D., Paterson, A., Tingey, S., and Hawk, J. (1994). Generation-means analysis and quantitative trait locus mapping

- of anthracnose stalk rot genes in maize. *Theor. Appl. Genet.* 89, 413–418. doi: 10.1007/BF00225375
- Kang, H., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548
- Khan, S. N. (2007). *Macrophomina phaseolina* as causal agent for charcoal rot of sunflower. *Mycopathologia* 5, 111–118. Available online at: <http://111.68.103.26/journals/index.php/mycopath/article/viewFile/218/115>
- Khokhar, M. K., Hooda, K. S., Sharma, S. S., and Singh, V. (2014). Post flowering stalk rot complex of maize - present status and future prospects. *Maydica* 59, 226–242. Available online at: <https://journals-crea.4science.it/index.php/maydica/article/view/1003/850>
- Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29. doi: 10.1186/1746-4811-9-29
- Kotchoni, S. O., Kuhns, C., Ditzer, A., Kirch, H.-H., and Bartels, D. (2006). Over-expression of different aldehyde dehydrogenase genes in *Arabidopsis thaliana* confers tolerance to abiotic stress and protects plants against lipid peroxidation and oxidative stress. *Plant Cell Environ.* 29, 1033–1048. doi: 10.1111/j.1365-3040.2005.01458.x
- Krishna, K. M., Kulakarni, N., and Kumar, R. S. (2013). Genetic analysis of post flowering stalk rot resistance in maize (*Zea mays* L.). *Indian J. Genet.* 73, 328–331. doi: 10.5958/j.0975-6906.73.3.049
- Kulich, I., Pečenková, T., Sekereš, J., Smetana, O., Fendrych, M., Foissner, I., et al. (2013). Arabidopsis exocyst subcomplex containing subunit EXO70B1 is involved in autophagy-related transport to the vacuole. *Traffic* 14, 1155–1165. doi: 10.1111/tra.12101
- Kumar, M., Lal, H. C., and Jha, M. M. (1996). Assessment of yield loss due to post-flowering stalk rot in maize. *J. Appl. Biol.* 8, 90–92.
- Lal, S., and Singh, I. S. (1984). Breeding for resistance to downy mildews and stalk rots in maize. *Theor. Appl. Genet.* 69, 111–119. doi: 10.1007/BF00272879
- Lannoo, N., and Van Damme, E. J. (2014). Lectin domains at the frontiers of plant defense. *Front. Plant Sci.* 5:397. doi: 10.3389/fpls.2014.00397
- Liao, C., Peng, Y., Ma, W., Liu, R., Li, C., and Li, X. (2012). Proteomic analysis revealed nitrogen-mediated metabolic, developmental, and hormonal regulation of maize (*Zea mays* L.) ear growth. *J. Exp. Bot.* 63, 5275–5288. doi: 10.1093/jxb/ers187
- Ma, C., Ma, X., Yao, L., Liu, Y., Du, F., Yang, X., et al. (2017). qRfg3, a novel quantitative resistance locus against gibberella stalk rot in maize. *Theor. Appl. Genet.* 130, 1723–1734. doi: 10.1007/s00122-017-2921-5
- Mahmoud, A. F., Abou-Elwafa, S. F., and Shehzad, T. (2018). Identification of charcoal rot resistance QTLs in sorghum using association and *in silico* analyses. *J. Appl. Genet.* 59, 243–251. doi: 10.1007/s13353-018-0446-5
- Maize AICRP. (2014). *Annual Progress Report Rabi*, eds O. P. Yadav, B. Kumar, G. K. Chikkappa, J. Kaul, A. K. Singh, C. M. Parihar, et al. New Delhi: Indian Institution of Maize Research.
- Mao, H., Wang, H., Liu, S., Li, Z., Yang, X., Yan, J., et al. (2015). A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. *Nat. Commun.* 6:8326. doi: 10.1038/ncomms9326
- Minic, Z., Rihouey, C., Do, C. T., Lerouge, P., and Jouanin, L. (2004). Purification and characterization of enzymes exhibiting β -D-xylosidase activities in stem tissues of *Arabidopsis*. *Plant Physiology* 135, 867–878. doi: 10.1104/pp.104.041269
- Mir, Z. R., Singh, P. K., Zaidi, P. H., Vinayan, M. T., Krishna, M. K., Vemula, A. K., et al. (2018). Genetic analysis of resistance to post flowering stalk rot in tropical germplasm of maize (*Zea mays* L.). *Crop Protect.* 106, 42–49. doi: 10.1016/j.cropro.2017.12.004
- Morant, M., Bak, S., Möller, B. L., and Werck-Reichhart, D. (2003). Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. *Curr. Opin. Biotechnol.* 14, 151–162. doi: 10.1016/S0958-1669(03)00024-7
- Payak, M. M., and Sharma, R. C. (1983). “Disease rating scales in maize in India,” in *Techniques of Scoring for Resistance to Important Diseases of Maize*. New Delhi: All India Coordinated Maize Improvement Project; Indian Agricultural Research Institute, 1–4.
- Pè, M. E., Gianfranceschi, L., Taramino, G., Tarchini, R., Angelini, P., Dani, M., et al. (1993). Mapping quantitative trait loci (QTLs) for resistance to *Gibberella zeae* infection in maize. *Molec. Gen. Genet.* 241, 11–16. doi: 10.1007/BF00280195
- Prasanna, B. M. (2018). “Maize in Asia – status, challenges and opportunities,” in *Book of Extended Summaries, 13th Asian Maize Conference and Expert Consultation on Maize for Food, Feed, Nutrition and Environmental Security*, eds B. M. Prasanna, A. Das, and K. K. Kaimeniyi (Ludhiana: CIMMYT), 1–8.
- Prasanna, B. M., Cairns, J. E., Zaidi, P. H., Beyene, Y., Gowda, M., Makumbi, D., et al. (2021). Beat the stress: breeding for climate resilience in maize for the tropical rainfed environments. *Theor. Appl. Genet.* 134, 1729–1752. doi: 10.1007/s00122-021-03773-7
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org>.
- Rafalski, J. A. (2002). Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Sci.* 162, 329–333. doi: 10.1016/S0168-9452(01)00587-8
- Rafalski, J. A. (2010). Association genetics in crop improvement. *Curr. Opin. Plant Biol.* 13, 174–180. doi: 10.1016/j.pbi.2009.12.004
- Rashid, Z., Singh, P. K., Vemuri, H., Zaidi, P. H., Prasanna, B. M., and Nair, S. K. (2018). Genome-wide association study in Asia-adapted tropical maize reveals novel and explored genomic regions for sorghum downy mildew resistance. *Sci. Rep.* 8:366. doi: 10.1038/s41598-017-18690-3
- Rashid, Z., Sofi, M., Harlapur, S. I., Kachapur, R. M., Dar, Z. A., Singh, P. K., et al. (2020). Genome-wide association studies in tropical maize germplasm reveal novel and known genomic regions for resistance to Northern corn leaf blight. *Sci. Rep.* 10:21949. doi: 10.1038/s41598-020-78928-5
- Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., et al. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14:R55. doi: 10.1186/gb-2013-14-6-r55
- Sakamoto, H., Maruyama, K., Sakuma, Y., Meshi, T., Iwabuchi, M., Shinozaki, K., et al. (2004). *Arabidopsis* Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiol.* 136:2734–2746. doi: 10.1104/pp.104.046599
- Schuler, M. A., and Werck-Reichhart, D. (2003). Functional genomics of P450s. *Ann. Rev. Plant Biol.* 54, 629–667. doi: 10.1146/annurev.arplant.54.031902.134840
- Sharma, R. C., De Leon, C., and Payak, M. M. (1993). Diseases of maize in South and South-East Asia: problems and progress. *Crop Prot.* 12, 414–422. doi: 10.1016/0261-2194(93)90002-Z
- Shiferaw, B., Prasanna, B. M., Hellin, J., and Banziger, M. (2011). Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Sec.* 3:307. doi: 10.1007/s12571-011-0140-5
- Singh, R. D. N., and Kaiser, S. A. K. M. (1991). Genetic analysis of resistance to charcoal rot of maize. *J. Mycopathol. Res.* 29, 103–109.
- Smith, G. S., and Wyllie, T. D. (1999). “Charcoal rot,” in *Compendium of Soybean Disease, 4th edn*, eds G. L. Hartman, J. B. Sinclair, and J. C. Rupe (St. Paul, MN: American Phytopathological Society), 29–31
- Smolarkiewicz, M., Skrzypczak, T., Michalak, M., Leśniewicz, K., Walker, J. R., Ingram, G., et al. (2014). Gamma-secretase subunits associate in intracellular membrane compartments in *Arabidopsis thaliana*. *J. Exp. Bot.* 65, 3015–3027. doi: 10.1093/jxb/eru147
- Swarts, K., Li, H., Navarro, A. R., An, D., Romay, M. C., Hearne, S., et al. (2014). Novel methods to optimize genotypic imputation for low-coverage, next-generation sequence data in crop plants. *Plant Genome* 7, 1–12. doi: 10.3835/plantgenome2014.05.0023
- Tesso, T., Ochanda, N., Claflin, L., and Tuinstra, M. (2009). An improved method for screening fusarium stalk rot resistance in grain sorghum [*Sorghum bicolor* (L.) Moench.]. *Afr. J. Plant Sci.* 3, 254–262. doi: 10.5897/AJPS.9000014
- Walter, S., Brennan, J. M., Arunachalam, C., Ansari, K. I., Hu, X., Khan, M. R., et al. (2008). Components of the gene network associated with genotype-dependent response of wheat to the *Fusarium mycotoxin* deoxynivalenol. *Funct. Integr. Genomics* 8, 421–427. doi: 10.1007/s10142-008-0089-4
- Walter, S., and Doohan, F. (2011). Transcript profiling of the phytotoxic response of wheat to the *Fusarium mycotoxin* deoxynivalenol. *Mycotox. Res.* 27, 221–230. doi: 10.1007/s12550-011-0099-2
- Wang, L., Zhang, Y., Zhu, X., Zhu, X., Li, D., Zhang, X., et al. (2017). Development of an SSR-based genetic map in sesame and identification of quantitative trait loci associated with charcoal rot resistance. *Sci. Rep.* 7:8349. doi: 10.1038/s41598-017-08858-2

- Wang, X., Liu, T., Li, C., and Zhao, Z. (2012). Gene expression profiles in maize (*Zea mays* L.) leaves inoculation with southern corn rust (*Puccinia polysora* Underw.). *Acta Physiol. Plant* 34, 997–1006. doi: 10.1007/s11738-011-0896-4
- Wen, Y., Wang, X., Xiao, S., and Wang, Y. (2012). Ectopic expression of VpALDH2B4, a novel aldehyde dehydrogenase gene from Chinese wild grapevine (*Vitis pseudoreticulata*), enhances resistance to mildew pathogens and salt stress in *Arabidopsis*. *Planta* 236, 525–539. doi: 10.1007/s00425-012-1624-z
- Wyllie, T. D. (1988). “Charcoal rot of soybeans — current status,” in *Soybean Diseases of the North Central Region*, eds T. D. Wyllie, and D. H. Scott (St. Paul, MN: APS Press), 106–113
- Yang, Q. (2010). A major QTL for resistance to Gibberella stalk rot in maize. *Theor. Appl. Genet.* 121, 673–687. doi: 10.1007/s00122-010-1339-0
- Zaidi, P. H., Nguyen, T., Ha, D. N., Thaitad, S., Ahmed, S., Arshad, M., et al. (2018). “Maize for changing climate - chasing the moving target,” in *Book of Extended Summaries, 13th Asian Maize Conference and Expert Consultation on Maize for Food, Feed, Nutrition and Environmental Security*, eds B. M. Prasanna, A. Das, and K. K. Kaimeniyi (Ludhiana: CIMMYT), 84–96.
- Zaidi, P. H., Seetharam, K., Krishna, G., Krishnamurthy, L., Gajanan, S., Babu, R., et al. (2016). Genomic regions associated with root traits under drought stress in tropical maize (*Zea mays* L.). *PLoS ONE* 11:e0164340. doi: 10.1371/journal.pone.0164340
- Zhang, D., Liu, Y., Guo, Y., Yang, Q., Ye, J., Chen, S., et al. (2012). Fine-mapping of qRfg2, a QTL for resistance to gibberella stalk rot in maize. *Theor. Appl. Genet.* 124, 585–596. doi: 10.1007/s00122-011-1731-4

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 Rashid, Kaur, Babu, Singh, Harlapur and Nair. 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:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Julio Huerta Espino,
Instituto Nacional de Investigaciones
Forestales, Agrícolas y Pecuarias
(INIFAP), Mexico
Dong Fang Ma,
Yangtze University, China

*Correspondence:

Houyang Kang
houyang.kang@sicau.edu.cn
Yunfeng Jiang
Jiangyunfeng@sicau.edu.cn
Guoyue Chen
gychen@sicau.edu.cn

[†]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: 10 August 2021

Accepted: 21 September 2021

Published: 12 October 2021

Citation:

Wang Y, Liang F, Guan F, Yao F,
Long L, Zhao X, Duan L, Wu Y, Li H,
Li W, Jiang Q, Wei Y, Ma J, Qi P,
Deng M, Zheng Y, Kang H,
Jiang Y and Chen G (2021) Molecular
Mapping and Analysis of an Excellent
Quantitative Trait Loci Conferring
Adult-Plant Resistance to Stripe Rust
in Chinese Wheat Landrace
Gaoxianguangtoumai.
Front. Plant Sci. 12:756557.
doi: 10.3389/fpls.2021.756557

Molecular Mapping and Analysis of an Excellent Quantitative Trait Loci Conferring Adult-Plant Resistance to Stripe Rust in Chinese Wheat Landrace Gaoxianguangtoumai

Yuqi Wang^{1,2†}, Fengying Liang^{1†}, Fangnian Guan¹, Fangjie Yao¹, Li Long¹, Xuyang Zhao¹, Luyao Duan¹, Yu Wu¹, Hao Li¹, Wei Li³, Qiantao Jiang^{1,2}, Yuming Wei^{1,2}, Jian Ma¹, Pengfei Qi¹, Mei Deng¹, Youliang Zheng^{1,2}, Houyang Kang^{1,2*}, Yunfeng Jiang^{1*} and Guoyue Chen^{1,2*}

¹Triticeae Research Institute, Sichuan Agricultural University, Chengdu, China, ²State Key Laboratory of Crop Gene Exploitation and Utilization in Southwest China, Sichuan Agricultural University, Chengdu, Sichuan, China, ³College of Agronomy, Sichuan Agricultural University, Chengdu, China

The Chinese wheat landrace “Gaoxianguangtoumai” (GX) has exhibited a high level of adult-plant resistance (APR) to stripe rust in the field for more than a decade. To reveal the genetic background for APR to stripe rust in GX, a set of 249 F_{6:8} (F₆, F₇, and F₈) recombinant inbred lines (RILs) was developed from a cross between GX and the susceptible cultivar “Taichung 29.” The parents and RILs were evaluated for disease severity at the adult-plant stage in the field by artificial inoculation with the currently predominant Chinese *Puccinia striiformis* f. sp. *tritici* races during three cropping seasons and genotyped using the Wheat 55K single-nucleotide polymorphism (SNP) array to construct a genetic map with 1,871 SNP markers finally. Two stable APR quantitative trait loci (QTL), QYr.GX-2AS and QYr.GX-7DS in GX, were detected on chromosomes 2AS and 7DS, which explained 15.5–27.0% and 11.5–13.5% of the total phenotypic variation, respectively. Compared with published Yr genes and QTL, QYr.GX-7DS and Yr18 may be the same, whereas QYr.GX-2AS is likely to be novel. Haplotype analysis revealed that QYr.GX-2AS is likely to be rare which presents in 5.3% of the 325 surveyed Chinese wheat landraces. By analyzing a heterogeneous inbred family (HIF) population from a residual heterozygous plant in an F₈ generation of RIL, QYr.GX-2AS was further flanked by KP2A_36.85 and KP2A_38.22 with a physical distance of about 1.37 Mb and co-segregated with the KP2A_37.09. Furthermore, three tightly linked Kompetitive allele-specific PCR (KASP) markers were highly polymorphic among 109 Chinese wheat cultivars. The results of this study can be used in wheat breeding for improving resistance to stripe rust.

Keywords: adult-plant resistance, QTL, stripe rust, Chinese wheat landrace, genetic mapping, heterogeneous inbred family

INTRODUCTION

Stripe rust (yellow rust), caused by *Puccinia striiformis* f. sp. *tritici* (*Pst*), is among the most harmful and widespread obligate pathogens of common wheat (*Triticum aestivum* L.) worldwide (Knott, 1989; Wellings, 2011). In China, stripe rust prevailed for several times in large wheat-growing areas and this caused serious yield losses (Zeng and Luo, 2006; Chen et al., 2014; Han and Kang, 2018). Since the 1950s, four severe epidemics of wheat stripe rust have occurred in China in 1950, 1964, 1990, and 2002, resulting in yield losses of 6.0, 3.2, 1.8, and 1.4 million tonnes, respectively (Li and Zeng, 2000; Wan et al., 2004). The main cause of the outbreaks is the emergence of new virulent races that overcome the widely deployed resistance genes (Chen and Kang, 2017). At present, new virulent *Pst* race CYR34 appears and overcomes the widely deployed *Yr* genes. In the meantime, the simplification of *Yr* genes in commercial wheat cultivars have not changed yet though a lot of cultivars are used in wheat production. The most efficient and economical method of controlling the disease is the use of genetic resistance (Liu et al., 2017; Wang et al., 2019). Continuous improvement in the resistance of wheat cultivars to cope with evolving races of *Pst* is a high priority to control stripe rust (Manickavelu et al., 2016).

To date, more than 300 genes or quantitative trait loci (QTL) for stripe rust resistance on the 21 wheat chromosomes have been reported (Rosewarne et al., 2013; McIntosh et al., 2019). In general, these resistance genes and QTL can be classified into two major classes: all-stage resistance (ASR) and adult-plant resistance (APR). ASR usually confers complete resistance during all growth stages and is simple to select during breeding. However, most ASR genes are race specific and encode nucleotide-binding and leucine-rich repeat (NLR) proteins, and therefore are effective against only a subset of *Pst* races. With regard to the dynamic rust pathogen populations of the virulent races, only a small number of the characterized ASR genes, such as *Yr5* (Marchal et al., 2018) and *Yr15* (Klymiuk et al., 2018), are still widely effective against currently dominant *Pst* race groups in China (Sharma-Poudyal et al., 2013; Wu et al., 2018).

In contrast, APR is effective starting at adult-plant growth stages and typically provides a degree of partial resistance. Although a few APR genes are race-specific (Milus et al., 2015), a greater proportion of APR genes including *Yr18* (Krattinger et al., 2009), *Yr29* (William et al., 2003), *Yr30* (Hayden et al., 2004), and *Yr46* (Moore et al., 2015) is non-race-specific and provides durable resistance to *Pst*. Of the three APR genes cloned to date, *Yr18* encodes a putative ATP-binding cassette transporter (Krattinger et al., 2009), *Yr36* encodes a kinase domain and a lipid-binding domain (Fu et al., 2009), and *Yr46* encodes a predicted hexose transporter (Moore et al., 2015). These genes represent different protein families compared with classical ASR genes (the NLR family) and provide unique mechanisms effective against a broader range of pathogens. As an example, *Yr18* has been globally used as a component of durable rust resistance in breeding programs and no evolution of increased virulence has been observed for almost 100 years (Krattinger et al., 2009). To achieve a high degree of durable

resistance, combining multiple APR genes into the same background has been considered as an important strategy for improvement of stripe rust resistance in wheat breeding.

Chinese wheat landraces are farmer-developed and maintained as traditional cultivars in China. These landraces harbor rich genetic diversity for stripe rust resistance. Numerous stripe rust genes or QTL have been identified, such as *Yr1* (Bansal et al., 2009), *Yr18* (Krattinger et al., 2009), *Yr81* (Gessese et al., 2019), *YrYL* (Wu et al., 2016a), *YrBai* (Ma et al., 2015), *Yrqb* (Cao et al., 2020), *QYr.caas-5AL* (Lan et al., 2010), *QYr.cau-6DL* (Zhang et al., 2017), *QYr.cau-2AL* (Wang et al., 2019a), *QYr.GTM-5DL* (Wu et al., 2020), and *QYr.AYH-5BL* (Long et al., 2021). Recently, our research program evaluated more than 1,000 Chinese wheat landrace accessions collected from all 10 agro-ecological zones (Zhou et al., 2017) for responses to stripe rust in the greenhouse and the field under inoculation with selected Chinese predominant races of *Pst* (Cheng et al., 2019; Long et al., 2019; Yao et al., 2019, 2020; Ye et al., 2019; Wang et al., 2021). Many resistant accessions of Chinese wheat landraces continually display APR to stripe rust in the field, providing a novel resistance resource for the breeding of wheat cultivars with durable resistance to stripe rust. Therefore, it is necessary and important to identify and develop new durable high-level APR resistance genes against stripe rust.

Gaoxianguangtoumai (GX) is a spring wheat landrace from Sichuan Province in southwest China, which is a regional center for overwintering and overwintering of the stripe rust pathogen. This landrace has exhibited a high degree of APR to stripe rust in the field for more than a decade, but little information is available on the genetic basis of resistance in this landrace. The objectives of the present study were to (1) identify the QTL conferring APR to stripe rust in a recombinant inbred line (RIL) population developed from the cross between GX and a susceptible cultivar, “Taichung 29” (TC 29, 2) validate and mendelize the novel QTL in a heterogeneous inbred family (HIF) population, and (3) develop tightly linked Kompetitive allele-specific PCR (KASP) markers for use in marker-assisted selection in breeding programs.

MATERIALS AND METHODS

Plant Materials and Races

The Chinese wheat landrace GX (accession number ZM7854 in National Germplasm Bank, China (NGBC) and AS1579 in Triticeae Research Institute, Sichuan Agricultural University) originating from Gao County, a county of Sichuan Province (28°26'N, 104°31'E). Because of high level of resistance to stripe rust for more than a decade, GX was crossed (as the female parent) with the highly stripe rust susceptible wheat cultivar TC 29. In total, 249 F₆s (F₆, F₇, and F₈) RILs derived from an individual F₁ plant were developed by single-seed descent. A KASP marker, KP2A_36.85 which was located around the peak of *QYr.GX-2AS*, was used to identify heterozygous lines from an F₈ generation of RIL. Through a single heterozygous plant was selected and selfed (Tuinstra et al., 1997), a HIF population of 130 individuals was generated for validating the

QYr.GX-2AS. The scheme for developing the genetic populations was showed in **Supplementary Figure S1**. A collection of 325 Chinese wheat landraces was genotyped with the 55K single-nucleotide polymorphism (SNP) array and further was used for marker haplotype analysis (Zhou et al., 2017). A panel of 109 Sichuan wheat cultivars was used to determine the polymorphism of markers tightly linked with QYr.GX-2AS. The highly stripe rust susceptible wheat cultivars “Mingxian 169,” “SY95-71,” and “Avocet S” (AvS) were used as susceptible controls in seedling and adult-plant tests throughout the study. Here, SY95-71 is a spring wheat line, selected from hexaploid triticale/wheat followed by backcrossing with wheat (Eronga 83/Fan6||Fan6; Shu et al., 1999). The line has been widely used in China as a highly susceptible stripe rust spreader genotype or susceptible control. The *Pst* races (comprising CYR32, CYR33, CYR34, G22-14, Su11-4, Su11-5, and Su11-7; Wu et al., 2016b; Huang et al., 2018) were kindly provided by the Plant Protection Institute of the Gansu Academy of Agricultural Sciences, Gansu, China.

Evaluation of Resistance to Stripe Rust

Seedling tests to evaluate the stripe rust resistance of GX and TC 29 were conducted in a greenhouse using two prevalent Chinese *Pst* races (CYR32 and CYR34). Five plants of each line were sown in a plastic pot filled with nutrient soil and grown in a controlled environment in the greenhouse. Seedlings were inoculated at the two-leaf stage with each *Pst* race in accordance with the protocol of Hickey et al. (2012). Inoculated plants were placed in a dew chamber at 10°C and 100% relative humidity for 24h in the dark, and then moved to separate growth chambers at 15–16°C with 12–14h of light daily. When the susceptible control “Mingxian 169” showed full sporulation, the infection type (IT) on the second leaf (approximately 15–18 days after inoculation) was scored using a 0–9 scale (Line and Qayoum, 1992). Plants with IT scores of 1 to 6 were considered resistant, whereas plants with IT scores of 7–9 were considered susceptible.

Assessments of adult-plant stripe rust responses were conducted at the Chongzhou Experimental Station (30°33'N, 103°39'E), Sichuan Agricultural University, Chengdu, China. The $F_{6,8}$ RILs population and the parental lines were evaluated for APR to stripe rust during the 2017–2018, 2018–2019, and 2019–2020 growing seasons (referred to as CZ2018, CZ2019, and CZ2020, respectively). The HIF population of 130 individuals was evaluated for APR to stripe rust during the 2020–2021. The phenotype data of HIF population were used for Chi-Squared analysis (3:1 ratio) and genetic mapping. In all tests, 20 seeds of each line were planted in rows 2m in length and spaced 30cm apart, with individual plants spaced 10cm apart. The susceptible cultivar TC 29 was planted in every 20th row as a susceptible control. To provide inoculum for infection, the susceptible cultivars SY95-71 and AvS were planted around the perimeter of the experimental area as spreaders. Artificial inoculation was conducted using a mixture of currently predominant *Pst* races in China (comprising CYR32, CYR33, CYR34, G22-14, Su11-4, Su11-5, and Su11-7). Stripe rust

response was first recorded by scoring the IT and disease severity (DS) when the susceptible checks SY95-71 and AvS showed more than 80% DS and was followed by two additional evaluations at 7 day intervals (i.e., three evaluations in total) for three randomly selected individual plants. The IT was recorded based on the 0–9 scale of Line and Qayoum (1992). The DS was scored as the percentage infected leaf area (0, 5, 10, 20, 40, 60, 80%, or 100%) in accordance with the Chinese National Standard, GB/T 15797-2011. The final DS (FDS) was used for phenotypic analysis.

Genotyping, Linkage Map Construction, and QTL Analysis

Genomic DNA was extracted from a single plant for each line of the wheat materials using the cetyltrimethylammonium bromide method (Stewart and Via, 1993). The two parents (GX and TC 29) and the 117 RILs were genotyped using the Axiom® Wheat 55K SNP array (53,036 markers) by the China Golden Marker Biotechnology Company Ltd. (Beijing, China). Monomorphic and SNP loci with a minor allele frequency less than 0.3 were excluded with further analysis (Ma et al., 2019). Polymorphic SNP markers were used to remove redundant markers in the binning step using the BIN function, with the parameters missing rate=20% and distortion value=0.01, implemented in QTL IciMapping v4.2 (Wang et al., 2019b). The binned markers were used for linkage map construction using the Kosambi mapping function (Kosambi, 1944) with JoinMap v4.0 (Van Ooijen, 2006). Mapping of QTL was performed using QTL IciMapping v4.2 based on inclusive composite interval mapping with the preset parameters Step=1 cM, value of *p* for entering variables (PIN)=0.001, and logarithm of the odds (LOD)=2.5.

To determine the effects of the QTL, the RILs were divided into four groups based on the presence/absence of the most closely linked flanking markers of QYr.GX-2AS and QYr.GX-7DS. In addition, the epistatic interactions between QYr.GX-2AS and QYr.GX-7DS were identified in RILs using QTL IciMapping v4.2 based on inclusive composite interval mapping of digenic epistatic QTL (ICIM-EPI) functionality with the preset parameters Step=1 cM, value of *p* for entering variables (PIN)=0.0001, and LOD=5.

Haplotype Analysis

Haplotype analysis was performed to identify haplotype variants for QYr.GX-2AS in a collection of 325 Chinese wheat landrace accessions (Zhou et al., 2017; Ye et al., 2019). The informative markers linked to QYr.GX-2AS were screened using the Wheat 55K or Wheat 660K SNP arrays in accordance with the method described by Long et al. (2021). The SNP genotype data and the phenotype data (FDS) were obtained from recently published studies (Cheng et al., 2019; Long et al., 2019; Yao et al., 2019, 2020; Ye et al., 2019; Wang et al., 2021). Haplotype variants were detected using Haploview v4.2.¹ The haplotypes detected in at least 10 accessions were considered to be major haplotypes.

¹<http://www.broad.mit.edu/mpg/haploview/>

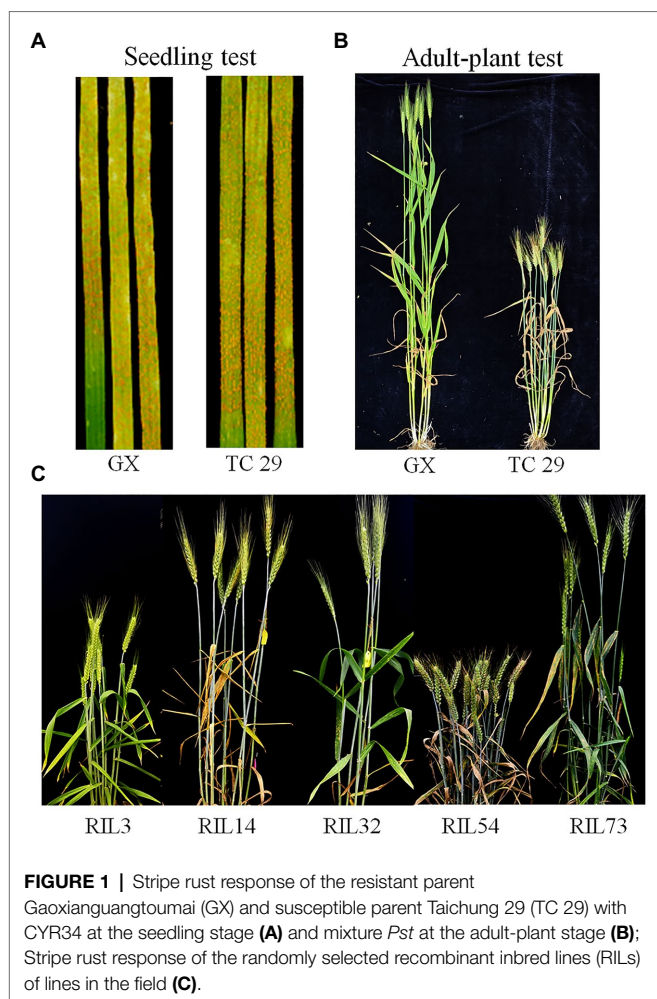


FIGURE 1 | Stripe rust response of the resistant parent Gaoxianguangtounmai (GX) and susceptible parent Taichung 29 (TC 29) with CYR34 at the seedling stage (A) and mixture *Pst* at the adult-plant stage (B); Stripe rust response of the randomly selected recombinant inbred lines (RILs) of lines in the field (C).

Boxplots were generated to display the average FDS of accessions carrying the different haplotypes. Haplotype data were combined with provenance information to examine the geographic distribution of the superior haplotypes in the 10 major agro-ecological production zones of Chinese wheat landraces.

Exome Capture Sequencing, Development of KASP Markers, and Genetic Mapping

Genomic DNA of the resistant parent GX was sequenced using the wheat exome capture sequencing protocol described by Dong et al. (2020). The raw sequence data have been submitted to GenBank under Bioproject no. PRJNA734801. The sequence variants were identified using the variant calling pipeline GATK4 (Heldenbrand et al., 2019). After QTL mapping, random SNPs in the target region to *QYr.GX-2AS* from the Wheat 55K array and exome capture sequencing were selected and converted to KASP markers using the PolyMarker online tool (Ramirez-Gonzalez et al., 2015). The specific KASP markers were used to screen the parents and a paired of NIL (selected from HIF population with a common genetic background but differing in *QYr.GX-2AS*) to confirm polymorphism before genotyping in the HIF population. The KASP assays were performed in

96-well format as 10 μ l reactions containing 2 μ l of 50–100 ng genomic DNA, 5 μ l of HiGeno 2 \times Probe Mix B, 0.24 μ M of each forward primer, 0.6 μ M of the common reverse primer, and double distilled water to make up the volume to 10 μ l. Each PCR was conducted using the BIO-RAD CFX96 qPCR system. Thermocycling was performed with a touchdown protocol: 95°C for 10 min; 95°C for 20 s and 61°C (−0.6°C per cycle) for 40 s for 10 cycles; and 95°C for 20 s and 55°C for 40 s for 38 cycles. Data analysis was performed manually using BIO-RAD CFX96 Manager 3.1.

The polymorphic KASP markers were used for validating the *QYr.GX-2AS* in the HIF population of 130 individuals. Linkage analysis was performed using JoinMap v4.0 (Kyazma BV, Wageningen, Netherlands; Van Ooijen, 2006) with a LOD threshold of 3.0. The Kosambi map function (Kosambi, 1944) was used to convert the recombination fractions to centi-Morgans. The linkage map was drawn using Mapdraw v2.1 (Liu and Meng, 2003). Three tightly linked markers for *QYr.GX-2AS* were further assessed in 109 wheat cultivars grown in Sichuan for checking the usefulness of the newly developed KASP markers for marker-assisted selected.

Data Analyses

Best linear unbiased prediction (BLUP) values for each RIL, ANOVA, Pearson's correlation coefficients, and broad-sense heritability (H^2) estimates were calculated using the "AOV" tool implemented in QTL IciMapping v4.2² (Wang et al., 2019b). The Chi-squared (χ^2) test with Excel 2016 was used to evaluate the goodness-of-fit for phenotype data of RILs population (1:1 ratio and 3:1 ratio) and HIF population (3:1 ratio). Student's *t*-tests ($p < 0.05$ and 0.01) were conducted with SPSS Statistics v17.0 (IBM Corp., Armonk, NY, United States) to evaluate the significance of differences between the two groups.

RESULTS

Stripe Rust Response of the Parents and RILs

Plants of GX were susceptible (IT=8) to CYR32 and CYR34 at the seedling stage (Figure 1A) but exhibited strong resistance (IT=3, FDS<10%) to mixed *Pst* races (comprising CYR32, CYR33, CYR34, G22-14, Su11-4, Su11-5, and Su11-7) at the adult-plant stage in three crop seasons from 2018 to 2020 (Figures 1B, 2; Supplementary Table S1), indicating that GX has effective APR to these prevalent Chinese *Pst* races. In all three environments, the average FDS of RILs for GX \times TC 29 was 12.5–15.7% in the field tests, and the distributions were skewed toward resistance (Figure 2). A total of 200 homozygous resistant lines (IT \leq 6) were consistently observed in the 249 RILs in all three field trials, and 142 lines of them showed high resistance similar to GX (IT \leq 3). In addition, 25 homozygous susceptible lines (IT \geq 7) were consistently observed in all three

²<http://www.isbreeding.net>

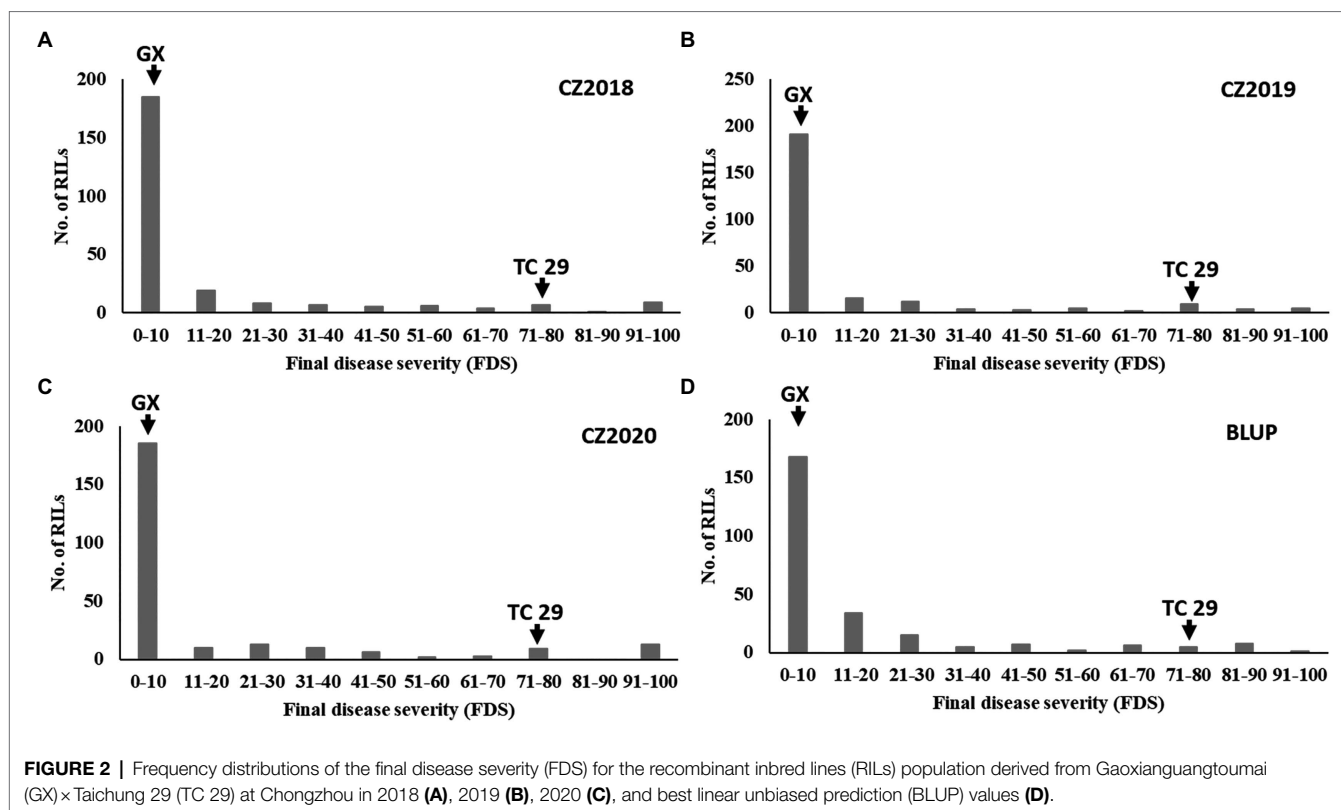


TABLE 1 | The summary of final disease severity (FDS) data for the recombinant inbred lines (RILs) population from the Gaoxianguangtoumai (GX) × Taichung 29 (TC 29) recorded in the fields at Chongzhou in 2018–2020.

| Environments | Parents | | RILs population | | | | |
|--------------|---------|-------|-----------------|------|------|-----|--------------------|
| | GX | TC 29 | Min-max | Mean | SD | CV | H ² (%) |
| CZ2018 (%) | 5.0 | 80.0 | 0–100 | 14.0 | 25.2 | 1.8 | 96.7 |
| CZ2019 (%) | 6.7 | 80.0 | 0–100 | 12.5 | 23.7 | 1.9 | |
| CZ2020 (%) | 5.0 | 80.0 | 0–100 | 15.7 | 27.0 | 1.7 | |
| BLUP (%) | 7.1 | 73.2 | 2.1–91.0 | 14.6 | 21.2 | 1.5 | |

SD, standard deviation; CV, coefficient of variation; and H², broad-sense heritability.

field trials. According to the homozygous phenotypes, the distribution of F_{68} families was not fit the expected ratios for a single gene (1:1 ratio; $\chi^2=136.11$, $p<0.001$) and two genes (3:1 ratio; $\chi^2=23.15$, $p<0.001$). The result indicated that the high level of resistance in GX was controlled by multiple genes (Figures 1C, 2; Supplementary Table S1). Broad-sense heritability (H^2) was 96.7% for FDS in all tests (Table 1). Correlation coefficients (R^2) for FDS of the RILs among the different environments were significant ($p<0.01$) and ranged from 0.82 to 0.95 (Supplementary Table S2).

Linkage Map Construction and QTL Analysis

A total of 1,871 markers were used to construct the linkage map which spanning a total length 2,799.12 cM for the GX × TC 29 population (Supplementary Table S3). The A, B, and D

genomes included 681 (36.40%), 669 (35.76%), and 521 (27.85%) markers covering lengths of 911.04, 855.71, and 1,032.37 cM with average marker intervals of 1.34, 1.28, and 1.98 cM, respectively (Supplementary Table S3).

Two high quality QTL, conferring APR to *Pst* races, was screened through further analysis (Table 2; Figures 3A,B). The most significant QTL, designated *QYr.GX-2AS*, was mapped to the short arm of chromosome 2AS and explained 15.5–27.0% phenotypic variation (Table 2; Figure 3A). The other QTL, designated *QYr.GX-7DS* and explaining 11.5–13.5% phenotypic variation, was located on the short arm of chromosome 7D where this gene overlaps with *Yr18* (Krattinger et al., 2009). The genetic distances analysis showed SNP markers *cssfr5* and *AX-110502471* flanking *QYr.GX-7DS* were 3.1 cM and 5.4 cM, respectively (Table 2; Figure 3B). Results indicated that it was highly likely that *QYr.GX-7DS* corresponded to *Yr18*.

TABLE 2 | Quantitative trait loci (QTL) for stripe rust resistance detected in the recombinant inbred lines (RILs) population from the Gaoxianguangtounmai (GX) × Taichung 29 (TC 29) using final disease severity (FDS) data across three environments and best linear unbiased prediction (BLUP) values.

| QTL | Environment | Trait | Chromosome | Left Marker | Right Marker | Chromosome interval (cM) | LOD | PVE (%) | Resistance source |
|------------|-------------|-------|------------|---------------|--------------|--------------------------|-----|---------|-------------------|
| QYr.GX-2AS | CZ2018 | FDS | 2AS | AX-109957471 | AX-110026721 | 2.8–3.1 | 8.1 | 27.0 | GX |
| | CZ2019 | | | AX-109957471 | AX-110026721 | 2.8–3.1 | 7.1 | 17.1 | |
| | CZ2020 | | | AX-109957471 | AX-110026721 | 2.8–3.1 | 5.2 | 15.5 | |
| | BLUP | | | AX-109957471 | AX-110026721 | 2.8–3.1 | 7.7 | 21.8 | |
| QYr.GX-7DS | CZ2018 | FDS | 7DS | cssfr5 (Yr18) | AX-110502471 | 93.9–102.4 | 3.4 | 11.6 | GX |
| | CZ2019 | | | cssfr5 (Yr18) | AX-110502471 | 93.9–102.4 | 3.2 | 11.5 | |
| | CZ2020 | | | cssfr5 (Yr18) | AX-110502471 | 93.9–102.4 | 3.6 | 12.4 | |
| | BLUP | | | cssfr5 (Yr18) | AX-110502471 | 93.9–102.4 | 4.0 | 13.5 | |

Clearly, the RILs that carried one of the QTL showed a lower FDS than those without any QTL (average FDS=63.4%; **Figure 3C**). The RILs carrying only QYr.GX-7DS showed 14.8% of the average FDS, whereas average FDS of lines with only QYr.GX-2AS was 9.3%. The lines with two QTL had the highest resistance level (average FDS=7.06%; **Figure 3C**), similar to that of GX. In addition, the epistatic interaction between QYr.GX-2AS and QYr.GX-7DS could be significantly detected in two field trials and Busing the ICIM-EPI functionality of the QTL IciMapping v4.2 (**Supplementary Table S4**). These results indicated that the high-level resistance in GX was contributed by these two QTL through additive and epistatic interactions, where QYr.GX-2AS provided relatively stronger resistance to *Pst* races than QYr.GX-7DS.

Haplotype Analysis of QYr.GX-2AS

To assess the distribution of QYr.GX-2AS among 325 Chinese wheat landraces, the favorable haplotype was identified by haplotype analysis and seven SNP markers tightly linked to QYr.GX-2AS were screened from the Wheat 55K or 660K SNP arrays (**Figures 4A–C**). Eight major haplotypes ($n > 10$) were detected in the panel (**Figures 4A,B**). GX and 15 other accessions clustered with Hap1 (**Supplementary Table S5**), which showed a frequency of about 5.3% in the total population (**Figure 4A**). Almost all accessions carrying Hap1, except one from Henan, were collected from Sichuan. The accessions carrying Hap1 showed 18.4% of the average FDS and thus were more strongly resistant to stripe rust than those accessions carrying other haplotypes (Hap2=37.2%, Hap3=24.1%, Hap4=47.7%, Hap5=21.5%, Hap6=39.0%, Hap7=27.0%, and Hap8=47.6%; **Figure 4C**). The above results suggested that Hap 1 was the favorable haplotype of QYr.GX-2AS and relatively rare in Chinese wheat landraces.

Validation and Mapping of QYr.GX-2AS

QYr.GX-2AS was further mapped finely using newly KASP markers developed from SNPs screened by exome capture

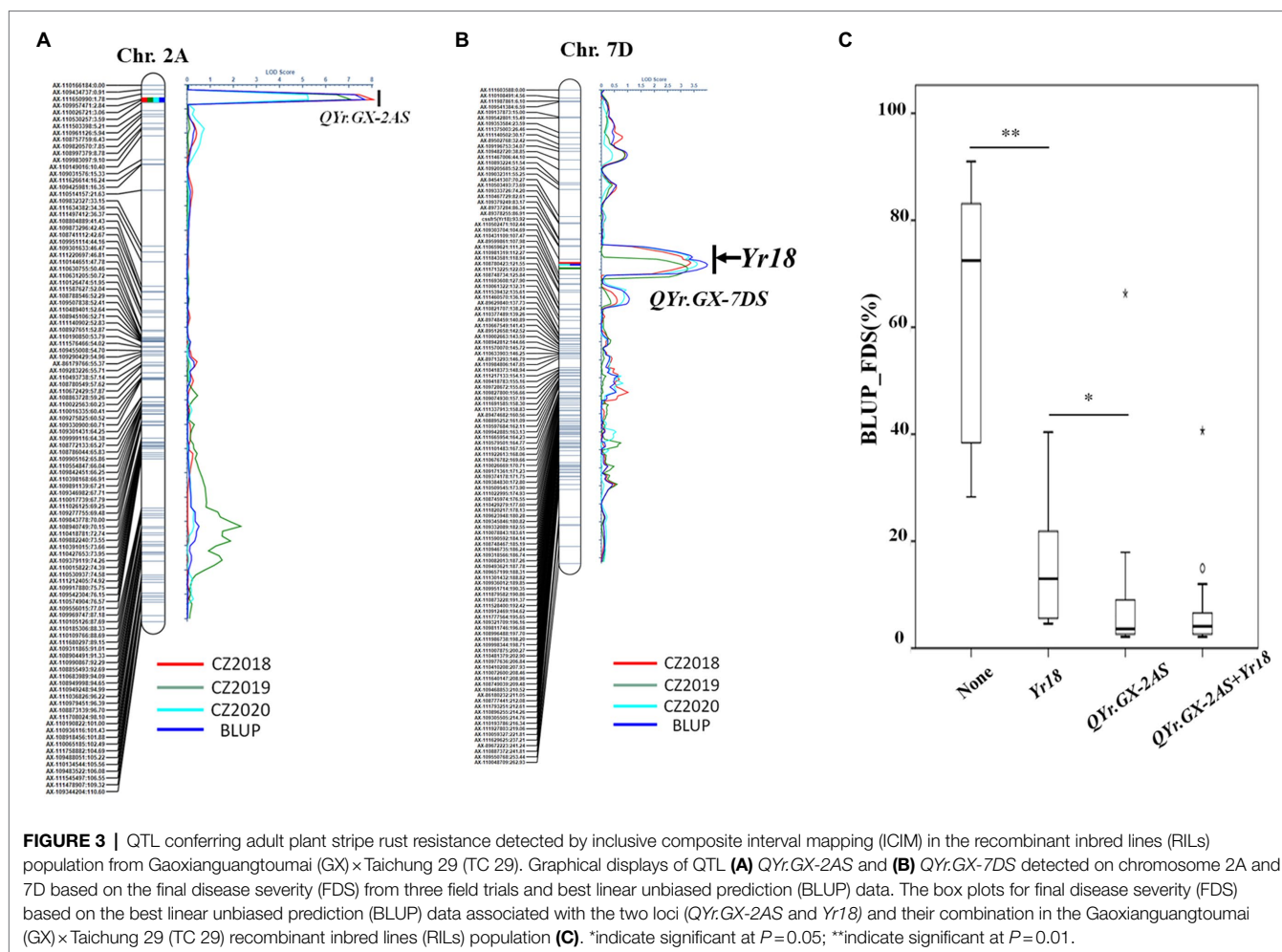
sequencing and the Wheat 55K array. Eleven markers were confirmed to be polymorphic between GX and TC 29 (**Supplementary Table S6**). Combined with the KASP marker KP2A_36.85 for QYr.GX-2AS and the marker *cssfr5* for Yr18, the HIF population of 130 individuals with a single locus QYr.GX-2AS was developed from a heterozygous plant (IT=4) in the F_8 generation of RILs (**Supplementary Figure S1**). No significant phenotypic differences were observed in the HIF population, except for APR to stripe rust (**Figure 5A**). With regard to stripe rust response in the field test, the HIF population could be clearly classifiable into 97 resistant (IT=3–4) and 33 susceptible (IT=8–9) individuals, which fits the expected ratio (3:1) for a single Mendelian factor (chi-square goodness-of-fit test, $\chi^2=0.01$, $p=0.92$; **Supplementary Table S7**). Using the newly developed 11 KASP markers (**Supplementary Table S7**) to construct the genetic map, QYr.GX-2AS was screened in 1.37 Mb interval between the KASP marker KP2A_36.85 and KP2A_38.22 and co-segregated with the KP2A_37.09 (**Figure 5B**).

Validation of KASP Markers for Marker-Assisted Selection

The molecular identification of 109 Chinese wheat cultivars was tested with three KASP markers KP2A_36.85 (G/A), KP2A_37.09 (A/C), and KP2A_38.22 (G/A; **Supplementary Figure S2; Supplementary Table S8**), which suggested that most of the cultivars could be amplified susceptible-specific alleles and showed 85.3, 99.1, and 95.4% polymorphism, respectively (**Supplementary Table S8**). Based on the above results, three KASP markers were valuable to apply QYr.GX-2AS in wheat breeding by marker-assisted selection.

DISCUSSION

It is the highest priorities to develop durable resistance to *Pst* races in wheat breeding during the past decade (Chen, 2013).



A large number of genes or QTL that confer various degrees of APR to stripe rust have been identified (Chen, 2013), but most only have minor effects on stripe rust response and are therefore difficult to use in breeding. Thus, the identification of new high quality *Yr* genes or QTL with APR is useful in wheat breeding. The Chinese wheat landrace GX has displayed a high degree of APR to stripe rust in the field for more than a decade in southwest China. Two QTLs conferring APR to *Pst* races tested were identified in GX, tentatively named as *QYr.GX-2AS* and *QYr.GX-7DS*, and mapped on chromosome 2AS and 7DS, respectively. In addition, the *QYr.GX-2AS* had a large effect in the reduction of stripe rust severity at adult-plant stages, which would be expected to have a great potential to pyramid this QTL with other *Yr* gene/QTL to develop wheat cultivars with high-level and durable resistance to *Pst* races.

QTL analysis is a useful procedure to reveal possible multiple loci when analyzing complex genetic traits, such as APR to stripe rust, in resistant germplasm. However, this procedure only allows approximate mapping of the QTL (Tanksley and Hewitt, 1988) owing to the heterogeneity in genetic backgrounds. The confidence interval of many QTL spans a considerable genetic distance and, as a result, molecular markers for these

QTL may not be reliably used in marker-assisted selection. As a strategy for accurate mapping of QTL in genetic analysis, HIF populations that allow the conversion of a quantitative trait into a Mendelian factor have been widely used for fine mapping and cloning of many important QTL in wheat, such as *Yr18* (Krattinger et al., 2009), *Yr36* (Fu et al., 2009), *Fhb1* (Su et al., 2019), and *Fhb7* (Wang et al., 2020). In the present research, a HIF population targeting *QYr.GX-2AS* was developed based on the method of heterogeneous inbred family analysis (Tuinstra et al., 1997). Members of this population were unambiguously classified as either resistant or susceptible and fitted the expected ratio (3:1) for a single Mendelian factor; thus, accurate mapping of the locus was possible. Analysis of the HIF population revealed that *QYr.GX-2AS*, flanked by *KP2A_36.85* and *KP2A_38.22*, was located in the interval 36.85 Mb to 38.22 Mb on chromosome 2AS. One KASP marker co-segregating with the targeted locus was successfully developed for marker-assisted selection.

Several genes that confer resistance to stripe rust have been identified on wheat chromosome 2AS, including *Yr17* (Bariana and McIntosh, 1993), *Yr56* (Bansal and Bariana 2014), *Yr69* (Hou et al., 2016), *Yr61* (Hao et al., 2011), and *YrSph* (Chen et al., 2012; Figure 6 and Supplementary Table S9). The genes *Yr17*,

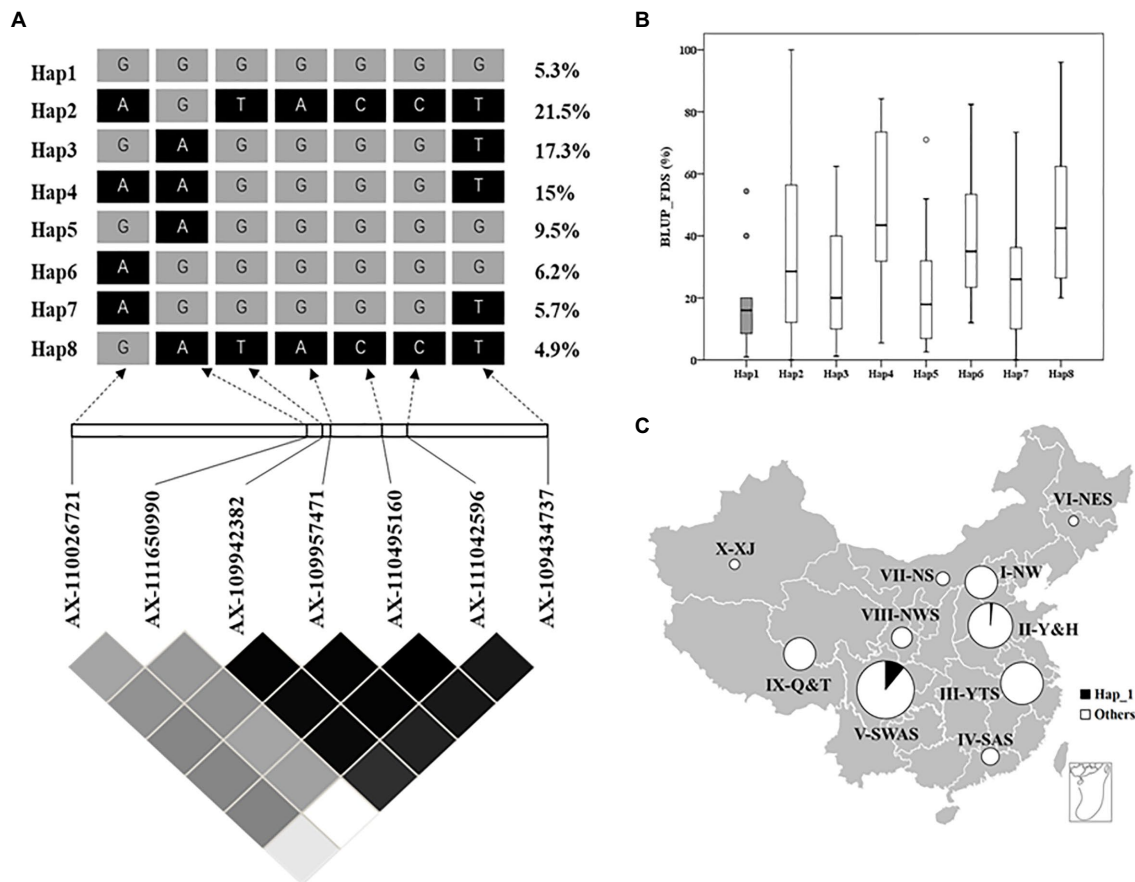
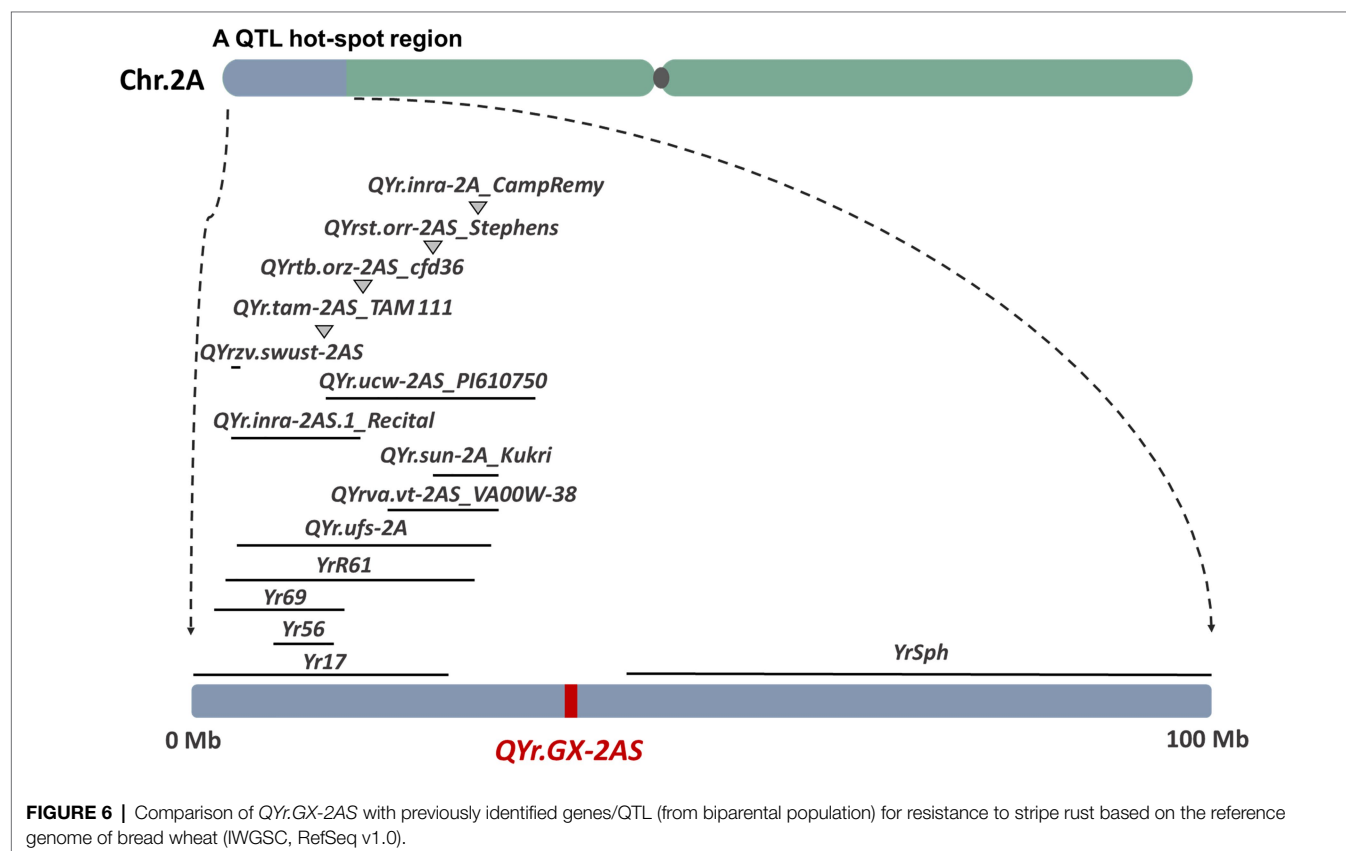
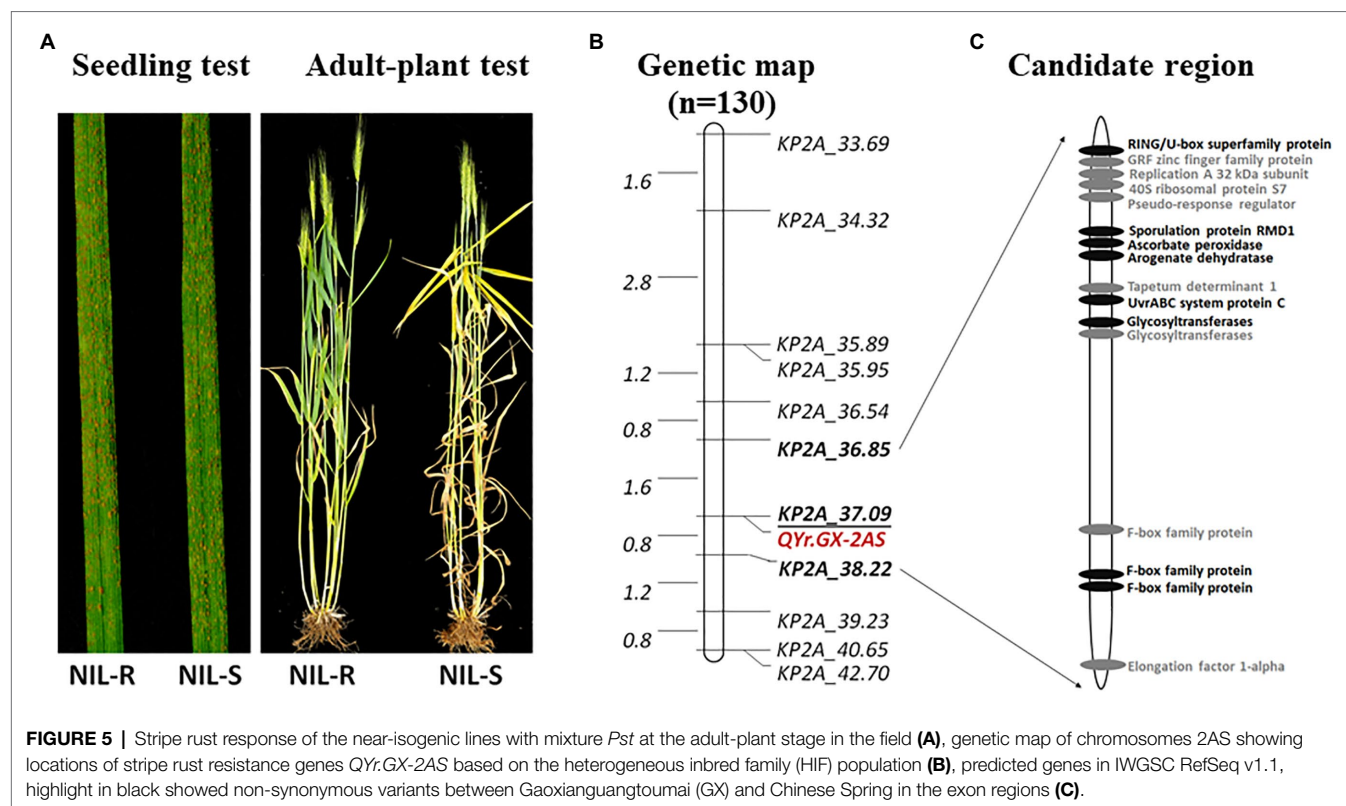


FIGURE 4 | Haplotype analysis of *QYr.GX-2AS* associated with stripe rust resistance in 325 Chinese wheat landraces. **(A)** LD heat map surrounding *QYr.GX-2AS*. The number on the right shows the distribution frequency of eight haplotypes in these Chinese wheat landraces. **(B)** Boxplot displays the mean final disease severity of the accessions carrying different haplotypes. **(C)** Frequencies of resistance allele of *QYr.GX-2AS* in Chinese wheat landraces in 10 major agro-ecological production zones of China.

Yr69, and *YrSph* confer ASR to stripe rust. Although recent studies suggest that *Yr17* also confers APR to stripe rust in the field, *QYr.GX-2AS* is likely to differ from *Yr17* because accessions of the Chinese wheat landrace GX that lack the 2N alien segment carry *Yr17*. *Yr56* is a major gene conferring APR to stripe rust that was identified in the Australian durum wheat cultivar “Wollaroi.” *Yr56* is flanked by *Xsun167* (*wPt-4,197*) and *Xsun168* (*wPt-9104*; Bansal and Bariana 2014), which corresponds to the “Chinese Spring” physical map region between 8.35 Mb and 14.28 Mb. *YrR61*, corresponding to the major-effect QTL *QYr.uga-2AS_26R61* conferring APR to stripe rust, was identified from the soft red winter wheat cultivar “Pioneer” and is flanked by the markers *Xbarc124* (3.78 Mb) and *Xgwm359* (28.20 Mb; Hao et al., 2011). Clearly, both *Yr56* and *YrR61* are located distant from *QYr.GX-2AS*. In addition, at least 20 QTL have been reported on chromosome 2AS, and most of them are located at a QTL hot-spot region in the distal end of 2AS (<30 Mb; **Figure 6**). For example, the QTL *QYr.tam-2AS_TAM 111* (Basnet et al., 2014) confer ASR to stripe rust. *QYr.ufs-2A* (Agenbag et al., 2012), *QYr.st.orr-2AS_Stephens* (Vazquez et al., 2012), and *QYr.sun-2A_Kukri* (Bariana et al., 2010) were all flanked by the basis

of a common DArT marker *XwPt-0003*, which were nearly with the *QYr.vt-2AS_VA00W-38* (Christopher et al., 2013) corresponds to the “Chinese Spring” physical map region 29.94 Mb. *QYr.tb.orz-2AS* (Vazquez et al., 2015) and *QYr.inra_2AS.1_Recital* (Dedryver et al., 2009) were located in 2AS close to marker *Xcfd36* (about 16.63 Mb) which are homologous to the *Yr17* introgression. The *QYr.ucw-2AS_PI610750* (Lowe et al., 2011), contributed by the synthetic derivative PI610750, is flanked by the *XwPt-3896* (13.14 Mb) and *Xwmc177* (33.70 Mb). *QYr.inra-2A_CampRemy* from Camp Remy (Mallard et al., 2005) is located by the *Xgwm382a* and *Xgwm359* (about 28.20 Mb). *QYr.zv.swust-2AS* (Zhou et al., 2021) flanked by *IWB7877* and *IWB72720* is derived from the wild emmer wheat (*T. dicoccoides*) accession Zavitan, corresponding to the “Chinese Spring” physical map region between 5.25 Mb and 5.33 Mb. Similarly, the other QTL identified by GWAS is located in different regions from *QYr.GX-2AS* on chromosome 2AS, expect for a minor locus *QYr.wsu-2A.1_IWA2526* (about 36.63 Mb). Hence, the large-effect QTL *QYr.GX-2AS* identified in the present study is unlikely to be the previously reported QTL. Anyway, the most powerful evidence still is gene sequencing on the target region after cloning *QYr.GX-2AS*.



According to gene annotation information in IWGSC RefSeq v1.1, 16 predicted genes are located in the candidate region for *QYr.GX-2AS* (Figure 5C; Supplementary Table S10). None of these genes is a classic NBS-LRR resistance gene. In addition, no annotations accorded with the protein types encoded by the APR genes *Yr18* (ABC transporter), *Yr36* (kinase-START), and *Yr48* (hexose transporter), implying that the candidate gene for *QYr.GX-2AS* might differ from known stripe rust resistance genes. Combined with exon sequencing data, eight predicted genes showed non-synonymous variants between GX and “Chinese Spring” in exon regions, including a RING/U-box, ascorbate peroxidase, glycosyltransferases, and F-box family protein, that may be involved in disease resistance. For confirmation of the candidate gene and cloning of *QYr.GX-2AS*, fine mapping to narrow the candidate interval will be performed using a large HIF population in future work.

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

YW and FL are responsible for the experiment, analyzed the data, and drafted the manuscript. FG, FY, LL, XZ, LD, YW, and HL carried out the phenotypic evaluation. WL, QJ, YW,

JM, PQ, MD, and YZ provided the resources and technique guidance. HK, YJ, and GC designed and carried out the experiment, formulated the questions, analyzed the data, and revised the manuscript. All authors have reviewed and approved the final manuscript.

FUNDING

This research was supported by the National Key Research and Development Program of China (2016YFD0100100), the International Science and Technology Cooperation and Exchanges Programs of Science and Technology Department of Sichuan Province (2019YFH0063), and the Applied Basic Research Programs of Sichuan Province (2021YJ0297).

ACKNOWLEDGMENTS

The authors are grateful to Prof. Qiu-Zhen Jia (Plant Protection Research Institute, Gansu Academy of Agricultural Sciences, Lanzhou, P. R. China) for providing the stripe rust races, and Prof. Li-Hui Li and Xiu-Quan Li (Chinese Academy of Agricultural Sciences) for plant materials (Sichuan wheat landraces) support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.756557/full#supplementary-material>

REFERENCES

- Agenbag, G. M., Pretorius, Z. A., Boyd, L. A., Bender, C. M., and Prins, R. (2012). Identification of adult plant resistance to stripe rust in the wheat cultivar cappelle-desprez. *Theor. Appl. Genet.* 125, 109–120. doi: 10.1007/s00122-012-1819-5
- Bansal, K., and Bariana, H. (2014). Mapping of stripe rust resistance gene *Yr56* in durum wheat cultivar Wollaroi. BGRI 2014 Technical workshop; March 22–25, 2014, Obregon, Mexico.
- Bansal, U. K., Hayden, M. J., Keller, B., Wellings, C. R., Park, R. F., and Bariana, H. S. (2009). Relationship between wheat rust resistance genes *Yr1* and *Sr48* and a microsatellite marker. *Plant Pathol.* 58, 1039–1043. doi: 10.1111/j.1365-3059.2009.02144.x
- Bariana, H. S., Bansal, U. K., Schmidt, A., Lehmsiek, A., Kaur, J., Miah, H., et al. (2010). Molecular mapping of adult plant stripe rust resistance in wheat and identification of pyramided QTL genotypes. *Euphytica* 176, 251–260. doi: 10.1007/s10681-010-0240-x
- Bariana, H. S., and McIntosh, R. A. (1993). Cytogenetic studies in wheat. XV. Location of rust resistance genes in VPM1 and their genetic linkage with other disease resistance genes in chromosome 2A. *Genome* 36, 476–482. doi: 10.1139/g93-065
- Basnet, B. R., Ibrahim, A. M. H., Chen, X. M., Singh, R., Mason, R. E., Bowden, R., et al. (2014). Molecular mapping of stripe rust resistance in hard red winter wheat TAM 111 adapted to the U.S. high plains. *Crop Sci.* 54, 1361–1373. doi: 10.2135/cropsci2013.09.0625
- Cao, J., Xu, Z., Fan, X., Zhou, Q., Ji, G., Wang, F., et al. (2020). Genetic mapping and utilization analysis of stripe rust resistance genes in a Tibetan wheat (*Triticum aestivum* L.) landrace Qubaichun. *Genet. Resour. Crop Evol.* 67, 1765–1775. doi: 10.1007/s10722-020-00938-z
- Chen, X. M. (2013). High-temperature adult-plant resistance, key for sustainable control of stripe rust. *Am. J. Plant Sci.* 4, 608–627. doi: 10.4236/ajps.2013.43080
- Chen, S. S., Chen, G. Y., Chen, H., Wei, Y. M., Li, W., Liu, Y. X., et al. (2012). Mapping stripe rust resistance gene *YrSph* derived from *Triticum sphaerococcum* Perc. with SSR, SRAP, and TRAP markers. *Euphytica* 185, 19–26. doi: 10.1007/s10681-011-0593-9
- Chen, X. M., and Kang, Z. S. (2017). “Introduction: history of research, symptoms, taxonomy of the pathogen, host range, distribution, and impact of stripe rust” in *Stripe Rust*. eds. X. M. Chen and Z. S. Kang (Berlin: Springer), 1–34.
- Chen, W. Q., Wellings, C., Chen, X. M., Kang, Z. S., and Liu, T. G. (2014). Wheat stripe (yellow) rust caused by *Puccinia striiformis* f. sp. *Tritici*. *Mol. Plant Pathol.* 15, 433–446. doi: 10.1111/mpp.12116
- Cheng, Y. K., Li, J., Yao, F. J., Long, L., Wang, Y. Q., Wu, Y., et al. (2019). Dissection of loci conferring resistance to stripe rust in Chinese wheat landraces from the middle and lower reaches of the Yangtze River via genome-wide association study. *Plant Sci.* 287:110204. doi: 10.1016/j.plantsci.2019.110204
- Christopher, M. D., Liu, S., Hall, M. D., Marshall, D. S., and Griffey, C. A. (2013). Identification and mapping of adult plant stripe rust resistance in soft red winter wheat VA00W-38. *Plant Breed.* 132, 53–60. doi: 10.1111/pbr.12015
- Dedryver, F., Paillard, S., Mallard, S., Robert, O., Trottet, M., Nègre, S., et al. (2009). Characterization of genetic components involved in durable resistance to stripe rust in the bread wheat “Renan”. *Phytopathology* 99, 968–973. doi: 10.1094/PHYTO-99-8-0968
- Dong, C. H., Zhang, L. C., Chen, Z. X., Xia, C., Gu, Y. Q., Wang, J. R., et al. (2020). Combining a new exome capture panel with an effective varBScore algorithm accelerates BSA-based gene cloning in wheat. *Front. Plant Sci.* 11:1249. doi: 10.3389/fpls.2020.01249

- Fu, D. L., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X. M., et al. (2009). A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science* 323, 1357–1360. doi: 10.1126/science.1166289
- Gessesse, M., Bariana, H., Wong, D., Hayden, M., and Bansal, U. (2019). Molecular mapping of stripe rust resistance gene *Yr81* in a common wheat landrace Aus27430. *Plant Dis.* 103, 1166–1171. doi: 10.1094/PDIS-06-18-1055-RE
- Han, D. J., and Kang, Z. S. (2018). Current status and future strategy in breeding wheat for resistance to stripe rust in China. *Plant Prot.* 44, 1–12. doi: 10.16688/j.zwbh.2018342
- Hao, Y., Chen, Z., Wang, Y., Bland, D., Buck, J., Brown-Guedira, G., et al. (2011). Characterization of a major QTL for adult plant resistance to stripe rust in US soft red winter wheat. *Theor. Appl. Genet.* 123, 1401–1411. doi: 10.1007/s00122-011-1675-8
- Hayden, M. J., Kuchel, H., and Chalmers, K. J. (2004). Sequence tagged microsatellites for the *Xgwm533* locus provide new diagnostic markers to select for the presence of stem rust resistance gene *Sr2* in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 1641–1647. doi: 10.1007/s00122-004-1787-5
- Heldenbrand, J. R., Baheti, S., Bockol, M. A., Drucker, T. M., Hart, S. N., Hudson, M. E., et al. (2019). Recommendations for performance optimizations when using GATK3.8 and GATK4. *BMC Bioinf.* 20:557. doi: 10.1186/s12859-019-3169-7
- Hickey, L., Wilkinson, P. M., Knight, C. R., Godwin, I. D., Kravchuk, O. Y., Aitken, E. A. B., et al. (2012). Rapid phenotyping for adult-plant resistance to stripe rust in wheat. *Plant Breed.* 131, 54–61. doi: 10.1111/j.1439-0523.2011.01925.x
- Hou, L. Y., Jia, J. Q., Zhang, X. J., Li, X., Yang, Z. J., Ma, J., et al. (2016). Molecular mapping of the stripe rust resistance gene *Yr69* on wheat chromosome 2AS. *Plant Dis.* 100, 1717–1724. doi: 10.1094/PDIS-05-15-0555-RE
- Huang, J., Jia, Q. Z., Zhang, B., Sun, Z. Y., Huang, M. M., and Jin, S. L. (2018). Epidemic forecasting of the new strains G22-9 (CYR34) and G22-14 of *Puccinia striiformis* f. sp. *tritici* in wheat in Gansu Province. *J. Plant Prot.* 45, 101–108. doi: 10.13802/j.cnki.zwbhxb.2018.2018910
- Klymiuk, V., Yaniv, E., Huang, L., Raats, D., Fatiukha, A., Chen, S. S., et al. (2018). Cloning of the wheat *Yr15* resistance gene sheds light on the plant tandem kinase-pseudokinase family. *Nat. Commun.* 9:3735. doi: 10.1038/s41467-018-06138-9
- Knott, D. R. (1989). The mode of inheritance of a type of dwarfism in common wheat. *Genome* 32, 932–933. doi: 10.1139/g89-533
- Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugenics* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x
- Krattinger, S. G., Lagudah, E. S., Spielmeier, W., Singh, R. P., Huerta-Espino, J., McFadden, H., et al. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323, 1360–1363. doi: 10.1126/science.1166453
- Lan, C. X., Liang, S. S., Zhou, X. C., Zhou, G., Lu, Q. L., Xia, X. C., et al. (2010). Identification of genomic regions controlling adult-plant stripe rust resistance in Chinese landrace Pingyuan 50 through bulked segregant analysis. *Phytopathology* 100, 313–318. doi: 10.1094/PHYTO-100-4-0313
- Li, Z. Q., and Zeng, S. M. (2000). 'Wheat Rusts in China'. China Agricultural Press, China
- Line, R. F., and Qayoum, A. (1992). *Virulence, Aggressiveness, Evolution and Distribution of Races of Puccinia Striiformis (the Cause of Stripe Rust of Wheat) in North America, 1968–87*. U.S. Department of Agriculture Technical Bulletin No. 1788 (Springfield: The National Technical Information Service), p44.
- Liu, B., Liu, T. G., Zhang, Z. Y., Jia, Q. Z., Wang, B. T., Gao, L., et al. (2017). Discovery and pathogenicity of CYR34, a new race of *Puccinia striiformis* f. sp. *tritici* in China. *Acta Phytopathol. Sin.* 47, 681–687. doi: 10.13926/j.cnki.apps.000071
- Liu, R. H., and Meng, J. L. (2003). MapDraw: a microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data. *Hereditas* 25, 317–321. doi: 10.3321/j.issn:0253-9772.2003.03.019
- Long, L., Yao, F. J., Guan, F. N., Cheng, Y. K., Duan, L. Y., Zhao, X. Y., et al. (2021). A stable QTL on chromosome 5BL combined with *Yr18* conferring high-level adult-plant resistance to stripe rust in Chinese wheat landrace Anyuehong. *Phytopathology*. doi: 10.1094/PHYTO-10-20-0465-R [Epub ahead of print]
- Long, L., Yao, F. J., Yu, C., Ye, X. L., Cheng, Y. K., Wang, Y. Q., et al. (2019). Genome-wide association study for adult-plant resistance to stripe rust in Chinese wheat landraces (*Triticum aestivum* L.) from the yellow and Huai river valleys. *Front. Plant Sci.* 10:596. doi: 10.3389/fpls.2019.00596
- Lowe, I., Cantu, D., and Dubcovsky, J. (2011). Durable resistance to the wheat rusts: integrating systems biology and traditional phenotype-based research methods to guide the deployment of resistance genes. *Euphytica* 179, 69–79. doi: 10.1007/s10681-010-0311-z
- Ma, D. F., Li, Q., Tang, M. S., Chao, K. X., Li, J. C., Wang, B. T., et al. (2015). Mapping of gene conferring adult plant resistance to stripe rust in Chinese wheat landrace Baidatou. *Mol. Breed.* 35, 157–165. doi: 10.1007/s11032-015-0244-2
- Ma, J., Qin, N. N., Cai, B., Chen, G. Y., Ding, P. Y., Zhang, H., et al. (2019). Identification and validation of a novel major QTL for all-stage stripe rust resistance on 1BL in the winter wheat line 20828. *Theor. Appl. Genet.* 132, 1363–1373. doi: 10.1007/s00122-019-03283-7
- Mallard, S., Gaudet, D., Aldeia, A., Abelard, C., Besnard, A. L., Sourdille, P., et al. (2005). Genetic analysis of durable resistance to yellow rust in bread wheat. *Theor. Appl. Genet.* 110, 1401–1409. doi: 10.1007/s00122-005-1954-3
- Manickavelu, A., Joukhadar, R., Jighly, A., Lan, C. X., Huerta-Espino, J., Stanikzai, A. S., et al. (2016). Genome wide association mapping of stripe rust resistance in afghan wheat landraces. *Plant Sci.* 252, 222–229. doi: 10.1016/j.plantsci.2016.07.018
- Marchal, C., Zhang, J., Zhang, P., Fenwick, P., Steuernagel, B., Adamski, N. M., et al. (2018). BED-domain-containing immune receptors confer diverse resistance spectra to yellow rust. *Nat. Plants* 4, 662–668. doi: 10.1038/s41477-018-0236-4
- McIntosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, J., Morris, C., Appels, R., et al. (2019). Catalogue of Gene Symbols for Wheat. Available at: <https://shigen.nig.ac.jp/wheat/komugi/genes/symbolClassList.jsp>
- Milus, E. A., Moon, D. E., Lee, K. D., and Mason, R. E. (2015). Race-specific adult-plant resistance in winter wheat to stripe rust and characterization of pathogen virulence patterns. *Phytopathology* 105, 1114–1122. doi: 10.1094/PHYTO-11-14-0305-R
- Moore, J. W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Aylife, M., Huerta-Espino, J., et al. (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 47, 1494–1498. doi: 10.1038/ng.3439
- Ramirez-Gonzalez, R. H., Uauy, C., and Caccamo, M. (2015). PolyMarker: a fast polyploid primer design pipeline. *Bioinform* 31, 2038–2039. doi: 10.1093/bioinformatics/btv069
- Rosewarne, G. M., Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., Lan, C. X., and He, Z. H. (2013). Quantitative trait loci of stripe rust resistance in wheat. *Theor. Appl. Genet.* 126, 2427–2449. doi: 10.1007/s00122-013-2159-9
- Sharma-Poudyal, D., Chen, X. M., Wan, A. M., Zhan, G. M., Kang, Z. S., Cao, S. Q., et al. (2013). Virulence characterization of international collections of the wheat stripe rust pathogen, *Puccinia striiformis* f. sp. *tritici*. *Plant Dis.* 97, 379–386. doi: 10.1094/PDIS-01-12-0078-RE
- Shu, H. L., Yang, Z. J., and Li, G. R. (1999). Selection and evaluation of a wheat line SY95 -71 as new yellow rust spreader. *J. Sichuan Agric. Univ.* 17, 12–16. doi: 10.16036/j.issn.1000-2650.1999.03.003
- Stewart, C. N., and Via, L. E. (1993). A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. *BioTechniques* 14, 748–751
- Su, Z., Bernardo, A., and Tian, B. (2019). A deletion mutation in TaHRC confers *Fhb1* resistance to Fusarium head blight in wheat. *Nat. Genet.* 51, 1099–1105. doi: 10.1038/s41588-019-0425-8
- Tanksley, S. D., and Hewitt, J. (1988). Use of molecular markers in breeding for soluble solids content in tomato — a re-examination. *Theor. Appl. Genet.* 75, 811–823. doi: 10.1007/BF00265610
- Tuinstra, M. R., Ejeta, G., and Goldsbrough, P. B. (1997). Heterogeneous inbred family (HIF) analysis: a method for developing near isogenic lines that differ at quantitative trait loci. *Theor. Appl. Genet.* 95, 1005–1011. doi: 10.1007/s001220050654
- Van Ooijen, J. W. (2006). *JoinMap 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations*. Wageningen: Kyazma BV.
- Vazquez, M. D., Peterson, C. J., Riera-Lizarazu, O., Chen, X., Heesacker, A., and Mundt, C. (2012). Genetic analysis of adult plant, quantitative resistance to stripe rust in wheat cultivar 'Stephens' in multi-environment trials. *Theor. Appl. Genet.* 124, 1–11. doi: 10.1007/s00122-011-1681-x

- Vazquez, M. D., Robert, Z., Peterson, C. J., Chen, X. M., Heesacker, A., and Mundt, C. C. (2015). Multi-location wheat stripe rust qtl analysis: genetic background and epistatic interactions. *Theor. Appl. Genet.* 128, 1307–1318. doi: 10.1007/s00122-015-2507-z
- Wan, A. M., Zhao, Z. H., Chen, X. M., He, Z. H., Jin, S. L., Jia, Q. Z., et al. (2004). Wheat stripe rust epidemic and virulence of *Puccinia striiformis* f. sp. *tritici* in China in 2002. *Plant Dis.* 88, 896–904. doi: 10.1094/PDIS.2004.88.8.896
- Wang, J. K., Li, H. H., Zhang, L. Y., and Meng, L. (2019b). *User's Manual of QTL Ici-Mapping ver. Quantitative Genetics Group, Institute of Crop Science, Chinese Academy of Agricultural Sciences CAAS/Genetic Resources Program*. International Maize and Wheat Improvement Center CIMMYT, Beijing/Mexico City.
- Wang, Z., Ren, J. D., Du, Z. Y., Che, M. Z., Zhang, Y. B., Quan, W., et al. (2019a). Identification of a major QTL on chromosome arm 2AL for reducing yellow rust severity from a Chinese wheat landrace with evidence for durable resistance. *Theor. Appl. Genet.* 132, 457–471. doi: 10.1007/s00122-018-3232-1
- Wang, H., Sun, S., and Ge, W. (2020). Horizontal gene transfer of *Fhb7* from fungus underlies Fusarium head blight resistance in wheat. *Science* 368:eaba5435. doi: 10.1126/science.aba5435
- Wang, Y. Q., Yu, C., Cheng, Y. K., Yao, F. J., Long, L., Wu, Y., et al. (2021). Genome-wide association mapping reveals potential novel loci controlling stripe rust resistance in a Chinese wheat landrace diversity panel from the southern autumn-sown spring wheat zone. *BMC Genomics* 22, 1–15. doi: 10.1186/s12864-020-07331-1
- Wang, L., Zheng, D., Zuo, S. X., Chen, X. M., Zhuang, H., Huang, L. L., et al. (2018). Inheritance and linkage of virulence genes in Chinese predominant race CYR32 of the wheat stripe rust pathogen *Puccinia striiformis* f. sp. *tritici*. *Front. Plant Sci.* 9:120. doi: 10.3389/fpls.2018.00120
- Wellings, C. R. (2011). Global status of stripe rust: A review of historical and current threats. *Euphytica* 179, 129–141. doi: 10.1007/s10681-011-0360-y
- William, M., Singh, R. P., Huerta-Espino, J., Ortiz-Islas, S., and Hoisington, D. (2003). Molecular marker mapping of leaf rust resistance gene *Lr46* and its association with stripe rust resistance gene *Yr29* in wheat. *Phytopathology* 93, 153–159. doi: 10.1094/PHYTO.2003.93.2.153
- Wu, J. H., Liu, S. J., Wang, Q. L., Zeng, Q. D., Mu, J. M., Huang, S., et al. (2018). Rapid identification of an adult plant stripe rust resistance gene in hexaploid wheat by high-throughput SNP array genotyping of pooled extremes. *Theor. Appl. Genet.* 131, 43–58. doi: 10.1007/s00122-017-2984-3
- Wu, J. H., Wang, Q. L., Chen, X. M., Wang, M. J., Mu, J. M., Lv, X. N., et al. (2016b). Stripe rust resistance in wheat breeding lines developed for Central Shaanxi, an overwintering region for *Puccinia striiformis* f. sp. *tritici* in China. *Can. J. Plant Pathol.* 38, 317–324. doi: 10.1080/07060661.2016.1206039
- Wu, X. L., Wang, J. W., Cheng, Y. K., Ye, X. L., Li, W., Pu, Z. E., et al. (2016a). Inheritance and molecular mapping of an all-stage stripe rust resistance gene derived from the Chinese common wheat land race 'Yilongtuomai'. *J. Hered.* 107, 463–470. doi: 10.1093/jhered/esw032
- Wu, Y., Wang, Y. Q., Yao, F. J., Long, L., Li, J., Li, H., et al. (2020). Molecular mapping of a novel QTL conferring adult plant resistance to stripe rust in Chinese wheat landrace 'Guangtoumai'. *Plant Dis.* 105:1187. doi: 10.1094/PDIS-07-20-1631-PDN
- Yao, F. J., Long, L., Wang, Y. Q., Duan, L. Y., Zhao, X. Y., Jiang, Y. F., et al. (2020). Population structure and genetic basis of the stripe rust resistance of 140 Chinese wheat landraces revealed by a genome-wide association study. *Plant Sci.* 301:110688. doi: 10.1016/j.plantsci.2020.110688
- Yao, F. J., Zhang, X. M., Ye, X. L., Li, J., Long, L., Yu, C., et al. (2019). Characterization of molecular diversity and genome-wide association study of stripe rust resistance at the adult plant stage in northern Chinese wheat landraces. *BMC Genet.* 20:38. doi: 10.1186/s12863-019-0736-x
- Ye, X. L., Li, J., Cheng, Y. K., Yao, F. J., Long, L., Yu, C., et al. (2019). Genome-wide association study of resistance to stripe rust (*Puccinia striiformis* f. sp. *tritici*) in Sichuan wheat. *BMC Plant Biol.* 19, 1–15. doi: 10.1186/s12870-019-1764-4
- Zeng, S. M., and Luo, Y. (2006). Long-distance spread and interregional epidemics of wheat stripe rust in China. *Plant Dis.* 90, 980–988. doi: 10.1094/PD-90-0980
- Zhang, H. Y., Wang, Z., Ren, J. D., Du, Z. Y., Quan, W., Zhang, Y. B., et al. (2017). A QTL with major effect on reducing stripe rust severity detected from a Chinese wheat landrace. *Plant Dis.* 101, 1533–1539. doi: 10.1094/PDIS-08-16-1131-RE
- Zhou, Y., Tang, H., Cheng, M. P., Dankwa, K., Chen, Z. X., Li, Z. Y., et al. (2017). Genome-wide association study for pre-harvest sprouting resistance in a large germplasm collection of Chinese wheat landraces. *Front. Plant Sci.* 8:401. doi: 10.3389/fpls.2017.00401
- Zhou, X. L., Zhong, X., Roter, J., Li, X., Yao, Q., Yan, J. H., et al. (2021). Genome-wide mapping of loci for adult-plant resistance to stripe rust in durum wheat Svevo using the 90K SNP array. *Plant Dis.* 105, 879–888. doi: 10.1094/PDIS-09-20-1933-RE

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 Wang, Liang, Guan, Yao, Long, Zhao, Duan, Wu, Li, Li, Jiang, Wei, Ma, Qi, Deng, Zheng, Kang, Jiang and Chen. 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.



Predicting Fusarium Head Blight Resistance for Advanced Trials in a Soft Red Winter Wheat Breeding Program With Genomic Selection

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Gurcharn Singh Brar,
University of British Columbia,
Canada

Yadong Huang,
University of Minnesota Twin Cities,
United States

*Correspondence:

Dylan L. Larkin
dylarkin92@gmail.com

†Present address:

Dylan L. Larkin,
Aardevo North America, Boise, ID,
United States
Richard Esten Mason,
Department of Soil and Crop
Sciences, Colorado State University,
Fort Collins, CO, United States
Amanda L. Holder,
Agriculture Department, Crowder
College, Neosho, MO, United States
Brian P. Ward,
Department of Horticulture and Crop
Science, Ohio State University,
Wooster, OH, United States

Specialty section:

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

Received: 26 May 2021

Accepted: 27 September 2021

Published: 22 October 2021

Citation:

Larkin DL, Mason RE, Moon DE,
Holder AL, Ward BP and
Brown-Guedira G (2021) Predicting
Fusarium Head Blight Resistance
for Advanced Trials in a Soft Red
Winter Wheat Breeding Program With
Genomic Selection.
Front. Plant Sci. 12:715314.
doi: 10.3389/fpls.2021.715314

**Dylan L. Larkin^{1*†}, Richard Esten Mason^{1†}, David E. Moon¹, Amanda L. Holder^{1†},
Brian P. Ward^{2†} and Gina Brown-Guedira^{2,3}**

¹ Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR, United States,

² USDA-ARS SEA, Plant Science Research, Raleigh, NC, United States, ³ Department of Crop and Soil Sciences, North Carolina State University, Raleigh, NC, United States

Many studies have evaluated the effectiveness of genomic selection (GS) using cross-validation within training populations; however, few have looked at its performance for forward prediction within a breeding program. The objectives for this study were to compare the performance of naïve GS (NGS) models without covariates and multi-trait GS (MTGS) models by predicting two years of F_{4:7} advanced breeding lines for three Fusarium head blight (FHB) resistance traits, deoxynivalenol (DON) accumulation, Fusarium damaged kernels (FDK), and severity (SEV) in soft red winter wheat and comparing predictions with phenotypic performance over two years of selection based on selection accuracy and response to selection. On average, for DON, the NGS model correctly selected 69.2% of elite genotypes, while the MTGS model correctly selected 70.1% of elite genotypes compared with 33.0% based on phenotypic selection from the advanced generation. During the 2018 breeding cycle, GS models had the greatest response to selection for DON, FDK, and SEV compared with phenotypic selection. The MTGS model performed better than NGS during the 2019 breeding cycle for all three traits, whereas NGS outperformed MTGS during the 2018 breeding cycle for all traits except for SEV. Overall, GS models were comparable, if not better than phenotypic selection for FHB resistance traits. This is particularly helpful when adverse environmental conditions prohibit accurate phenotyping. This study also shows that MTGS models can be effective for forward prediction when there are strong correlations between traits of interest and covariates in both training and validation populations.

Keywords: genomic selection, Fusarium head blight, wheat, resistance, multi-trait genomic selection, forward prediction

INTRODUCTION

Resistance to the disease Fusarium head blight (FHB) is important in wheat (*Triticum aestivum* L.) production, particularly in the Southeastern US. Fusarium head blight is a fungal disease caused by *Fusarium graminearum* and incurs nearly US\$4.2 billion in losses annually (Wilson et al., 2017). The *F. graminearum* pathogen produces the mycotoxin deoxynivalenol (DON), which is harmful for humans and animals that consume infected grain (FDA, 2010; Sobrova et al., 2010).

Traditionally, wheat breeders have primarily relied on phenotypic selection within their breeding programs to advance breeding material. However, phenotypic selection has its limitations, especially with low-heritability traits of interest that are difficult to phenotype. Difficulties with phenotyping are also compounded by genotype \times environment interactions that can lead to differential responses between genotypes across environments, reducing the accuracy of selections. Alternatives to phenotypic selection include marker assisted selection (MAS) and genomic selection (GS). Marker assisted selection can be effective for qualitative traits controlled by one or two genes or quantitative traits that are controlled by large-effect quantitative trait loci (QTL) (Xu and Crouch, 2008). However, MAS is less effective for complex quantitative traits controlled by many small-effect QTL (Bernardo and Yu, 2007; Heffner et al., 2009). Genomic selection is an effective alternative to both phenotypic selection and MAS, in that it incorporates allelic effects across the entire genome, making it ideal for quantitative traits. Genomic selection can also reduce the time within a breeding cycle, as two rounds of GS can be performed compared to one cycle of phenotypic selection allowing for greater genetic gain over time (Bernardo and Yu, 2007; Heffner et al., 2009; Asoro et al., 2013; Rutkoski et al., 2015).

Genomic selection was first applied to animal breeding, particularly in the dairy industry, but it has since been adapted by plant breeders over the last decade (Meuwissen et al., 2001; Heffner et al., 2009). Genomic selection uses a training population (TP), a panel of lines that have been phenotyped for a trait of interest and genotyped using whole-genome sequencing, to train a genomic prediction model. The genomic prediction model then uses relatedness between all genotypes to obtain genome-estimated breeding values (GEBVs) for breeding lines, otherwise known as the validation population (VP), that have only been genotyped. The breeder can then make selections based on the GEBVs for a trait of interest (Meuwissen et al., 2001).

Most studies involving GS have focused on increasing prediction accuracy by manipulating the TP and subsequently evaluating model performance through cross-validation within the TP (Habier et al., 2007; Heffner et al., 2009; Jannink et al., 2010; Combs and Bernardo, 2013; Akdemir et al., 2015; Isidro et al., 2015; Larkin et al., 2019). Many have also investigated the genomic prediction model used for GS analysis (Heslot et al., 2012). While these methods are valuable, few have researched the effectiveness of applying GS in breeding programs for forward prediction of breeding lines (Bernardo, 2016). However, when investigated, many have seen mixed results regarding prediction accuracy of forward prediction, compared to cross-validated prediction accuracy within TPs (Asoro et al., 2013; Combs and Bernardo, 2013; Massman et al., 2013; Michel et al., 2017; Belamkar et al., 2018; Calvert et al., 2020). Additionally, there are few, if any, studies that focus on forward prediction for FHB resistance in wheat as opposed to grain yield (GY) (Michel et al., 2016, 2017; Belamkar et al., 2018; Calvert et al., 2020).

In an evaluation of GS in the Kansas State University wheat breeding program, GS was used to predict GY in a TP where the prediction accuracy was between $r = 0.31$ and $r = 0.47$.

However, when the TP was used for forward prediction, the highest prediction accuracy between the GEBVs for GY in the preliminary yield trials (PYTs) and the actual phenotypic results for GY was $r = -0.16$ (Calvert et al., 2020). This trend was also observed in an evaluation of the University of Nebraska wheat breeding program, where GY data from PYTs from three years were used to predict the performance of a fourth year. When no lines for the fourth year were included in the TP, prediction accuracies for GY were between $r = 0.22$ and $r = 0.26$. However, as more lines from the fourth year were included in the TP, the prediction accuracy of GY for the remaining lines in the fourth year increased to between $r = 0.37$ and $r = 0.52$, when 90% of the lines from the fourth year were included in the TP (Belamkar et al., 2018). Phenotypic selection and GS were also compared in terms of selection accuracy between the PYT and advanced yield trial generations. Genomic selection outperformed phenotypic selection during the 2012 and 2015 seasons, where Nebraska experienced severe drought and disease stress. Even still, prediction accuracies were low, indicating that prediction accuracy is not the best indicator of GS success for forward prediction (Belamkar et al., 2018). Another study using forward prediction for GY in wheat adapted to central Europe found that the use of GS ($r = 0.39$) to select high performing lines for multiple-environment trials was far better than phenotypic selection ($r = 0.21$) (Michel et al., 2017).

In addition to traditional GS, researchers have begun investigating the efficacy of multi-trait GS (MTGS). Multi-trait GS uses mixed models that incorporate secondary traits that are genetically correlated with a trait of interest as covariates to improve the prediction accuracy for the trait of interest (Calus and Veerkamp, 2011; Jia and Jannink, 2012; Covarrubias-Pazaran et al., 2018). Multi-trait GS can improve prediction accuracies for low-heritability traits when high-heritability secondary traits are used as covariates (Calus and Veerkamp, 2011; Guo et al., 2014; Jia et al., 2018). Many studies have evaluated MTGS models for cross-validation, particularly for GY in wheat using high-throughput phenotyping traits (Rutkoski et al., 2016; Sun et al., 2017; Crain et al., 2018; Lozada and Carter, 2019; Guo et al., 2020). Others have evaluated resistance traits related to FHB in wheat using phenological traits, such as heading date (HD) and plant height (PH), or other FHB resistance traits as covariates (Rutkoski et al., 2012; Schulthess et al., 2018; Steiner et al., 2019; Larkin et al., 2020; Moreno-Amores et al., 2020). Few have evaluated the use of MTGS for forward prediction. One study used high-throughput phenotyping traits as a covariate in a MTGS model for forward prediction of GY in wheat, though the prediction accuracy was unfavorable unless a large TP was used (Calvert et al., 2020). Therefore, our aim is to validate the use of MTGS models compared to naïve GS (NGS) models to predict FHB resistance in wheat, using secondary FHB resistance traits regularly collected throughout the season within a breeding program based on results from Larkin et al. (2020).

The University of Arkansas soft red winter wheat (SRWW) breeding program makes over 800 unique crosses per year. Progenies are then tested over the following 10 seasons prior to releasing a new cultivar (Mason et al., 2018). Breeding

lines are not evaluated for FHB resistance traits until the F_{4:7} advanced (ADV) and F_{4:8} elite (ARE) trials, where they are evaluated in misted and inoculated FHB disease nurseries at two locations in a RCBD design with two replications. Selections are made based on three FHB resistance traits: type II resistance, which is resistance to the spread of FHB within a spike, otherwise known as severity (SEV) (Schroeder and Christensen, 1963); type III resistance, or resistance to Fusarium damaged kernels (FDK) (Argyris et al., 2003; Goral et al., 2019); and type IV resistance, or resistance to DON accumulation (Mesterhazy, 1995).

Some elite lines are also grown in regional statewide variety testing trials, as well as the USDA-ARS Uniform Eastern (UE) and Southern nurseries (US), Southeastern University Grains (Sungrains) cooperative nurseries, and foundation seed increases. The UE and US nurseries include approximately 36 elite breeding lines from public and private SRWW breeding programs in the Southern and Eastern US, grown between 22 and 36 locations with between one and three replications per location annually (Boyles et al., 2019). The Sungrains cooperative consists of Southeastern US SRWW breeding programs that performs regional testing within the Southeastern US (Harrison et al., 2017; Johnson et al., 2017; Mason et al., 2018; Boyles et al., 2019). Select breeding lines from the ADV and ARE are grown in these regional Sungrains nurseries.

In theory, GS can improve selection accuracy in the early generations of the breeding program for FHB resistance traits while also reducing time and resources spent for phenotyping. In this study, we evaluated the selection accuracy of GS from the advanced through elite generations and compared to phenotypic selection through forward prediction using NGS and MTGS models. The three goals for this study were to: (1) compare NGS and MTGS with phenotypic selection for three FHB resistance traits, including DON, FDK, and SEV for new breeding lines that have not been phenotyped at the advanced generations; (2) compare the selection accuracy between NGS, MTGS, and phenotypic selection between the advanced and elite generations of the University of Arkansas SRWW breeding program; and (3) compare the response to selection between NGS, MTGS, and phenotypic selection between the advanced and elite generations of the University of Arkansas SRWW breeding program.

MATERIALS AND METHODS

Plant Materials

Breeding Materials

Two generations of the ADV trials, 2017–2018 and 2018–2019, consisting of F_{4:7} breeding lines from the University of Arkansas wheat breeding program and doubled haploid (DH) lines developed through the Sungrains cooperative, were used as VPs to predict three FHB traits, DON, FDK, and SEV. Approximately 20% of breeding lines from the ADV18 and ADV19 yield trials were selected and advanced to the ARE19 and ARE20 yield trials for the 2018–2019 and 2019–2020 growing seasons, respectively. Genotypes were advanced based on both GS and phenotypic selection (Table 1).

Training Populations

A population of 355 SRWW genotypes was used as the initial 2018 TP (TP18_FHB) for this study to predict GEBVs for DON, FDK, and SEV in the ADV18 trial. The population consisted of 187 genotypes from the University of Arkansas, 87 from Louisiana State University, 40 from North Carolina State University, 38 from the University of Georgia, and one genotype each from Syngenta AG, Pioneer Hi-Bred International, Inc., and Virginia Polytechnic Institute and State University (Larkin et al., 2020). The 2019 TP (TP19_FHB) for the three FHB traits consisted of the 355 genotypes from TP18_FHB, as well as the 104 genotypes from the ADV18 trial.

Experimental Design and Trait Measurements

Winter wheat is planted during the fall and harvested during the late spring in the southern United States, therefore the growing season spans two years. The TP18_FHB genotypes were evaluated for three FHB resistance traits, including DON, FDK, and SEV, over four seasons between 2014 and 2017 at two locations, at the Milo J. Shult Agricultural Research and Extension Center in Fayetteville, AR, United States (FAY) and the Newport Research and Extension Center near Newport, AR, USA (NPT). The data collection and experimental design methods were outlined in Larkin et al. (2020), as TP18_FHB was the same population used in their study.

The AVD18, ADV19, and ARE19 FHB nurseries for the 2017–2018 and 2018–2019 growing seasons were grown at two locations, FAY and NPT, in a randomized complete block design (RCBD) with two replications per location using the same methods described with respect to the TP18_FHB and TP19_FHB populations in Larkin et al. (2020). This was also the case for the ARE20 FHB nursery; however, it was only grown in NPT during the 2019–2020 season due to poor growing conditions in FAY. Data were also collected for HD, PH, DON, FDK, and SEV for the FHB nurseries using methods described in Larkin et al. (2020).

Phenotypic Data Analyses

Phenotypic data was analyzed using a single stage mixed linear model within the PROC MIXED procedure in SAS 9.4 to obtain adjusted means for HD, PH, DON, FDK, and SEV (SAS Institute Inc., Cary, NC, United States). The following model was fit to the phenotypic data:

$$y_{ijk} = \mu + \text{genotype}_i + \text{rep}(\text{env})_{jk} + \text{env}_k + (\text{genotype} \times \text{env})_{ik} + \varepsilon_{ijk}$$

where y_{ijk} is the observed phenotype, μ is the population mean, genotype_i is the fixed effect of the i^{th} genotype, $\text{rep}(\text{env})_{jk}$ is the random effect of the j^{th} replication nested within the k^{th} location-year (or location) (env), env_k is the random effect of the k^{th} location-year (or location), $(\text{genotype} \times \text{env})_{ik}$ is the random effect of the interaction between genotype and location-year (or location), and ε_{ijk} is the residual error term, where $\varepsilon_{ijk} \sim N(0, I\sigma^2_\varepsilon)$, where I is an identity matrix and σ^2_ε is the residual error variance.

TABLE 1 | Description of the number of genotypes, composition, and experimental design of two generations of F_{4:7} advanced nurseries (ADV), and F_{4:8} elite nurseries (ARE), as well as the initial training population (TP18_FHB) used to predict three Fusarium head blight (FHB) resistance traits, including deoxynivalenol (DON) accumulation, Fusarium damaged kernels (FDK), and severity (SEV).

| Trial ^a | Generation ^b | Conventional lines | DH lines | Total | Location(s) | Rep(s) | Design ^c |
|--------------------|-------------------------|--------------------|----------|-------|-------------|--------|---------------------|
| TP18_FHB | – | 355 | – | 355 | 9 | 2 | RCBD |
| ADV18 | F4:7/DH | 64 | 40 | 104 | 2 | 2 | RCBD |
| ADV19 | F4:7/DH | 50 | 70 | 120 | 2 | 2 | RCBD |
| ARE19 | F4:8/DH | 16 | 6 | 22 | 2 | 2 | RCBD |
| ARE20 | F4:8/DH | 12 | 11 | 23 | 1 | 2 | RCBD |

^aTrial types and the years each were grown. TP18_FHB was grown over four years between 2013–2014 and 2016–2017; 18, 2017–2018; 19, 2018–2019; 20, 2019–2020.

^bBreeding trials consisted of conventionally bred genotypes as well as doubled haploid (DH) genotypes.

^cRCBD, randomized complete block design.

Phenotypic Pearson correlations were calculated between DON, FDK, HD, PH, and SEV within TP18_FHB and TP19_FHB as well as the ADV and ARE FHB nurseries using the multivariate function in JMP Pro 15.2.0 software (SAS Institute Inc., Cary, NC). Entry mean-based broad-sense heritability (H^2) was calculated for each trait using the following equation:

$$H^2 = \frac{\sigma_{\text{genotype}}^2}{\sigma_{\text{genotype}}^2 + \frac{\sigma_{\text{genotype} \times \text{env}}^2}{n_{\text{env}}} + \frac{\sigma_{\varepsilon}^2}{n_{\text{env}} \times n_{\text{rep}}}}$$

where $\sigma_{\text{genotype}}^2$ is the genotypic variance, $\sigma_{\text{genotype} \times \text{env}}^2$ is the variance of the interaction between genotype and location-year, n_{env} is the number of location-years where the trait was evaluated, σ_{ε}^2 is the residual error variance, and n_{rep} is the number of replications within each location-year. Variance components were obtained from the single stage mixed linear model described above for each trait using the PROC MIXED procedure in SAS 9.4. Narrow-sense heritability (h^2) was calculated using the “marker_h²” function within the “heritability” package in R v4.0.3 software for TP19_FHB due to a lack of shared genotypes within the TP (Kruijer et al., 2015; R Core Team, 2020). The analysis used a genome relationship matrix obtained from the “A.mat” function within the “rrBLUP” package in Rv4.0.3 software using the marker set described below as well as the abovementioned phenotypic data (Endelman, 2011; Endelman and Jannink, 2012; R Core Team, 2020).

Genotyping by Sequencing

All genotypes were genotyped using genotyping by sequencing (GBS) using methods described in Larkin et al. (2020). Single nucleotide polymorphism (SNP) calling was performed using the TASSEL 5.0 GBSv2 pipeline using 64 base tag length and a minimum tag count of five (Bradbury et al., 2007). Reads were aligned to the International Wheat Genome Sequencing Consortium (IWGSC) RefSeq v1.0 “Chinese Spring” wheat reference sequence (Appels et al., 2018) using the Burrows-Wheeler aligner version 0.7.17 (Li and Durbin, 2009).

Raw SNP data generated from the TASSEL pipeline were filtered using PLINK software (Purcell et al., 2007) to remove taxa with more than 85% missing data and heterozygosity greater than 30%. Genotypic data were then filtered to select

for biallelic SNPs with minor allelic frequency of greater than five percent, less than 20% missing data, and heterozygosity less than or equal to 10%. Missing marker data were then imputed using BEAGLE software, based on windows encompassing the entire chromosome (Browning et al., 2018). Markers were again filtered after imputation to select SNP markers with minor allele frequency greater than five percent and heterozygosity of less than equal to 10% using PLINK software. Markers aligning to unassembled contigs were also removed for a final genotypic dataset of 5,202 SNP markers.

Principal component analyses were performed within each of the TPs to evaluate the genetic relationships between subpopulations using the PCA function in TASSEL 5.0. These relationships between the first three principal components were visualized for each TP using the “scatterplot3d” package in R v4.0.3 software (Ligges et al., 2018; R Core Team, 2020).

Genomic Selection

Two different models were tested for both TPs to obtain GEBVs for DON, FDK, and SEV for the ADV18 and ADV19 trials. The first model was a naïve genomic BLUP (GBLUP) model with no covariates (NGS). The second model was a MTGS GBLUP model where DON was predicted using FDK and HD as covariates, FDK was predicted using DON and SEV as covariates, and SEV was predicted using FDK and PH as covariates. The optimal covariate combinations for the MTGS models were determined in Larkin et al. (2020) for the FHB traits.

Cross Validation

Mean prediction accuracies between the NGS and MTGS models for each TP were obtained using a five-fold cross-validation analysis performed using the Genomic Selection function in TASSEL 5.0 (Bradbury et al., 2007). The GBLUP model used for the analyses is described as follows:

$$y = X\beta + Zu + \varepsilon_i$$

where u is a vector of genotype effects, which is assumed to have a normal distribution $u \sim N(0, G\sigma_u^2)$, where G is the genomic relationship matrix, obtained using the Kinship function within TASSEL 5.0, which uses the same methodology as the “rrBLUP” package in R (Endelman, 2011; Endelman and Jannink, 2012),

and σ_u^2 is the variance of the individual genotype effects; β is a vector of fixed effects; X is a design matrix relating fixed effects to phenotypic observations (y); Z is a design matrix relating random effects to phenotypic observations; and ε_i is the residual error at the i th locus, which is assumed to have a normal distribution $\varepsilon_i \sim N(0, I\sigma_\varepsilon^2)$, where I is the identity matrix and σ_ε^2 is the residual error variance. The GEBV from the GBLUP model is equivalent to the sum of all allele effects of a genotype from the ridge regression BLUP (RR-BLUP) model (VanRaden, 2008; Endelman, 2011).

The five-fold cross-validation approach randomly divided the TP into five equal sized groups. Four of the five groups were then used as the TP to train the GBLUP model to calculate GEBVs for the fifth group, serving as the VP, where the phenotypic values were set as missing. In the case of the MTGS models, the phenotypic data for the covariate traits were used as a fixed effect in the model. The GEBVs calculated for the VP were compared to the actual phenotypic values using a Pearson correlation. The five-fold cross-validation process was repeated over 100 iterations for a total of 500 iterations. The mean prediction accuracies between the NGS and MTGS models were compared between both TPs using a generalized linear mixed model (GLMM) and Fisher's LSD with an α of 0.05, implemented in PROC GLIMMIX in SAS 9.4. Mean prediction accuracy comparisons between the NGS and MTGS models for each TP were visualized using the "yarr" package in R v4.0.3 (Phillips, 2017; R Core Team, 2020).

Forward Prediction

Both TPs were then used to obtain predictions for their respective VPs using the NGS and MTGS GBLUP models associated with each trait. For example, TP18_FHB was used to calculate GEBVs for DON, FDK, and SEV for the ADV18 trial using the NGS and MTGS models (Table 2).

Once GEBVs for each trait for each model were obtained, GEBVs were compared to the adjusted mean of the trait of interest for each genotype in the following generation using a Pearson correlation using the multivariate function in JMP 15.2.0 software. For example, GEBVs for DON obtained for ADV18 were compared to the adjusted mean DON for each genotype across the ADV18 and ARE19 generations. This serves as a form of prediction accuracy for the respective model and TP. A scatterplot visualizing the comparison between GEBVs and adjusted means across years for each genotype, as well as individual genotypes advanced to the next generation, was created using the "ggplot2" package in R v4.0.3 for each model for each TP (Wickham et al., 2016; R Core Team, 2020). Selection accuracy was also determined as the percentage of genotypes advanced to the ARE generation that were above average based on GEBVs from the NGS or MTGS models as well as above average based on phenotypic values.

Response to selection was also compared between the NGS and MTGS models and phenotypic selection, based on the adjusted means from the ADV generations for FHB traits, using a selection pressure of 50%. The response to selection formula is as follows:

$$R = H^2S$$

where H^2 was the broad-sense heritability calculated as above, and S is the selection differential, calculated as $S = \mu_{\text{Selected}} - \mu_{\text{Unselected}}$ where μ_{Selected} is the mean of the phenotypic data for the top 50% of genotypes selected for genotypes in the ARE generations using either phenotypic selection, NGS, or MTGS, and $\mu_{\text{Unselected}}$ is the mean of the full unselected population of the genotypes in the ARE generation of the breeding cycle (Falconer and McKay, 1996; Arruda et al., 2016b; Lozada et al., 2020).

RESULTS

Variation in Fusarium Head Blight Resistance Traits

Both FHB TPs as well as the ADV and ARE FHB trials had significant variation for all five traits. The ADV18 FHB trial had the highest mean DON and FDK, but it also had the lowest mean SEV. The ARE20 FHB trial had the lowest mean DON and FDK, likely due to stronger genetic resistance (Table 2). All trials also had significant correlations between the three FHB traits. Correlations between DON and HD were consistently positive, however, the correlations were not significant with smaller population sizes, while DON was significantly correlated with PH only in ADV19. There were generally negative correlations between FDK and PH apart from ADV19, however, the significance of the correlations between FDK and PH were not significant with smaller population sizes. There were strong negative correlations between SEV and HD and PH for nearly all trials, however, they were not significant for smaller populations. High heritability was also observed for all three FHB traits in addition to HD and PH (Table 2).

Population Structure

Genotyping by sequencing identified 5,202 SNPs across the entire wheat genome after filtering and imputation. The number of SNP markers were unevenly distributed between genomes, where the B genome had the largest number of markers (2,315), followed by the A (2,210) and D (677) genomes, which was consistent with other studies using GBS SNPs (Arruda et al., 2016a; Larkin et al., 2020). The chromosome with the largest number of SNPs was 3B at 477, while the chromosome with the smallest number was 4D (38). The proportion of heterozygosity within the dataset was 2.5% and the average minor allele frequency was 21.6%.

The PCA of the initial TP18_FHB population showed two primary clusters within the population. Genotypes from all breeding programs appeared in both clusters, although there was evidence of sub-clustering by breeding program within the two main clusters. This clustering has also been observed in other studies using SRWW populations adapted to the Southeastern US and is hypothesized to result from the large number of linked SNPs called between lines with and without a translocation from *Triticum timopheevii* Zhuk., which harbors stem rust (*Puccinia graminis* f. sp. *tritici*) and powdery mildew (*Blumeria graminis* f. sp. *tritici*) resistance genes *Sr36* and *Pm6* (Nyquist, 1962; Benson et al., 2012; Sarinelli et al., 2019; Larkin et al., 2020). The population structure was generally low, where the first three

TABLE 2 | Descriptive statistics, Pearson phenotypic correlations, and heritabilities (H^2) for adjusted means for two training populations, two advanced $F_{4,7}$ nurseries, and two elite $F_{4,8}$ nurseries for three Fusarium head blight (FHB) resistance traits, including deoxynivalenol (DON), Fusarium damaged kernels (FDK), and severity (SEV) as well as heading date (HD) and plant height (PH).

| Trial ^a | Trait | Summary statistics | | | | | | Correlations | | | |
|--------------------|-----------------|--------------------|--------|--------|--------|-------|----------|---------------------|----------------------|---------------------|-----------------|
| | | Mean | Min | Max | Range | SD | H^{2b} | DON ^c | FDK | SEV | HD ^d |
| FHB_TP18 | DON | 10.53 | 0.08 | 92.80 | 92.72 | 11.35 | 0.74 | — | — | — | — |
| | FDK | 32.49 | 0.00 | 100.00 | 100.00 | 29.93 | 0.79 | 0.40*** | — | — | — |
| | SEV | 27.88 | 0.00 | 100.00 | 100.00 | 25.78 | 0.82 | 0.32*** | 0.73*** | — | — |
| | HD | 94.69 | 74.00 | 118.00 | 44.00 | 10.19 | 0.90 | 0.25*** | −0.05 ^{ns†} | −0.10* | — |
| | PH ^e | 90.27 | 56.46 | 121.87 | 65.40 | 10.06 | 0.91 | 0.01 ^{ns} | −0.29*** | −0.36*** | 0.34*** |
| FHB_TP19 | DON | 14.26 | 6.15 | 37.50 | 31.35 | 4.59 | 0.60 | — | — | — | — |
| | FDK | 38.22 | 6.00 | 92.12 | 86.12 | 14.86 | 0.68 | 0.45*** | — | — | — |
| | SEV | 28.67 | 3.75 | 91.71 | 87.96 | 12.97 | 0.93 | 0.12* | 0.55*** | — | — |
| | HD | 97.91 | 86.76 | 116.50 | 29.74 | 8.30 | 0.92 | 0.31*** | 0.02 ^{ns} | −0.54*** | — |
| | PH | 90.40 | 71.12 | 113.03 | 41.91 | 6.96 | 0.74 | 0.00 ^{ns} | −0.29*** | −0.31*** | 0.16*** |
| ADV18 | DON | 16.64 | 3.60 | 51.50 | 47.90 | 7.60 | 0.62 | — | — | — | — |
| | FDK | 39.32 | 2.00 | 75.00 | 73.00 | 16.29 | 0.77 | 0.62*** | — | — | — |
| | SEV | 15.44 | 0.00 | 85.00 | 85.00 | 14.88 | 0.38 | 0.27** | 0.54*** | — | — |
| | HD | 112.22 | 108.00 | 117.00 | 9.00 | 2.16 | 0.90 | 0.28** | −0.10 ^{ns} | −0.34*** | — |
| | PH | 89.46 | 68.58 | 119.38 | 50.80 | 8.53 | 0.71 | −0.05 ^{ns} | −0.22* | −0.18* | 0.25*** |
| ADV19 | DON | 10.09 | 0.12 | 74.50 | 74.38 | 10.08 | 0.61 | — | — | — | — |
| | FDK | 31.01 | 0.00 | 98.00 | 98.00 | 23.94 | 0.83 | 0.86*** | — | — | — |
| | SEV | 25.60 | 0.00 | 95.00 | 95.00 | 25.54 | 0.45 | 0.76*** | 0.86*** | — | — |
| | HD | 102.38 | 97.00 | 109.00 | 12.00 | 2.35 | 0.81 | 0.10 ^{ns} | −0.04 ^{ns} | −0.17 ^{ns} | — |
| | PH | 81.20 | 63.50 | 101.60 | 38.10 | 7.16 | 0.74 | 0.29*** | 0.09 ^{ns} | 0.01 ^{ns} | 0.35*** |
| ARE19 | DON | 8.51 | 0.59 | 64.10 | 63.51 | 8.32 | 0.50 | — | — | — | — |
| | FDK | 27.04 | 1.00 | 95.00 | 94.00 | 21.06 | 0.71 | 0.84*** | — | — | — |
| | SEV | 23.55 | 0.00 | 90.00 | 90.00 | 22.87 | 0.43 | 0.74*** | 0.86*** | — | — |
| | HD | 102.20 | 98.00 | 108.00 | 10.00 | 2.27 | 0.84 | 0.01 ^{ns} | −0.28 ^{ns} | −0.28 ^{ns} | — |
| | PH | 80.06 | 53.34 | 93.98 | 40.64 | 7.14 | 0.76 | 0.21 ^{ns} | −0.12 ^{ns} | −0.18 ^{ns} | 0.41* |
| ARE20 | DON | 7.30 | 0.99 | 19.30 | 18.31 | 3.95 | 0.78 | — | — | — | — |
| | FDK | 15.21 | 2.00 | 60.00 | 58.00 | 12.49 | 0.84 | 0.78*** | — | — | — |
| | SEV | 16.83 | 0.00 | 60.00 | 60.00 | 13.11 | 0.76 | 0.65*** | 0.82*** | — | — |
| | HD | 99.49 | 94.00 | 111.00 | 17.00 | 3.39 | 0.76 | 0.09 ^{ns†} | 0.09 ^{ns} | −0.08 ^{ns} | — |
| | PH | 89.69 | 76.20 | 101.60 | 25.40 | 6.20 | 0.87 | 0.01 ^{ns} | −0.05 ^{ns} | −0.11 ^{ns} | 0.48** |

^aTP, training population; ADV, $F_{4,7}$ advanced FHB trial; ARE, $F_{4,8}$ elite FHB trial.

^bBroad-sense heritability for FHB_TP18, ADV18, ADV19, ARE19, and ARE20 calculated using entry-mean based heritability. Narrow-sense heritability was calculated for FHB_TP19.

^cDON was recorded in $\mu\text{g g}^{-1}$, whereas FDK and SEV were recorded in percentage.

^dHeading date was recorded as day of year after 1st of January, when 50% of the heads were emerged from the flag leaf.

^ePlant height was recorded in inches from the surface of the soil to the tip of the head minus awns if present, but reported in centimeters here.

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

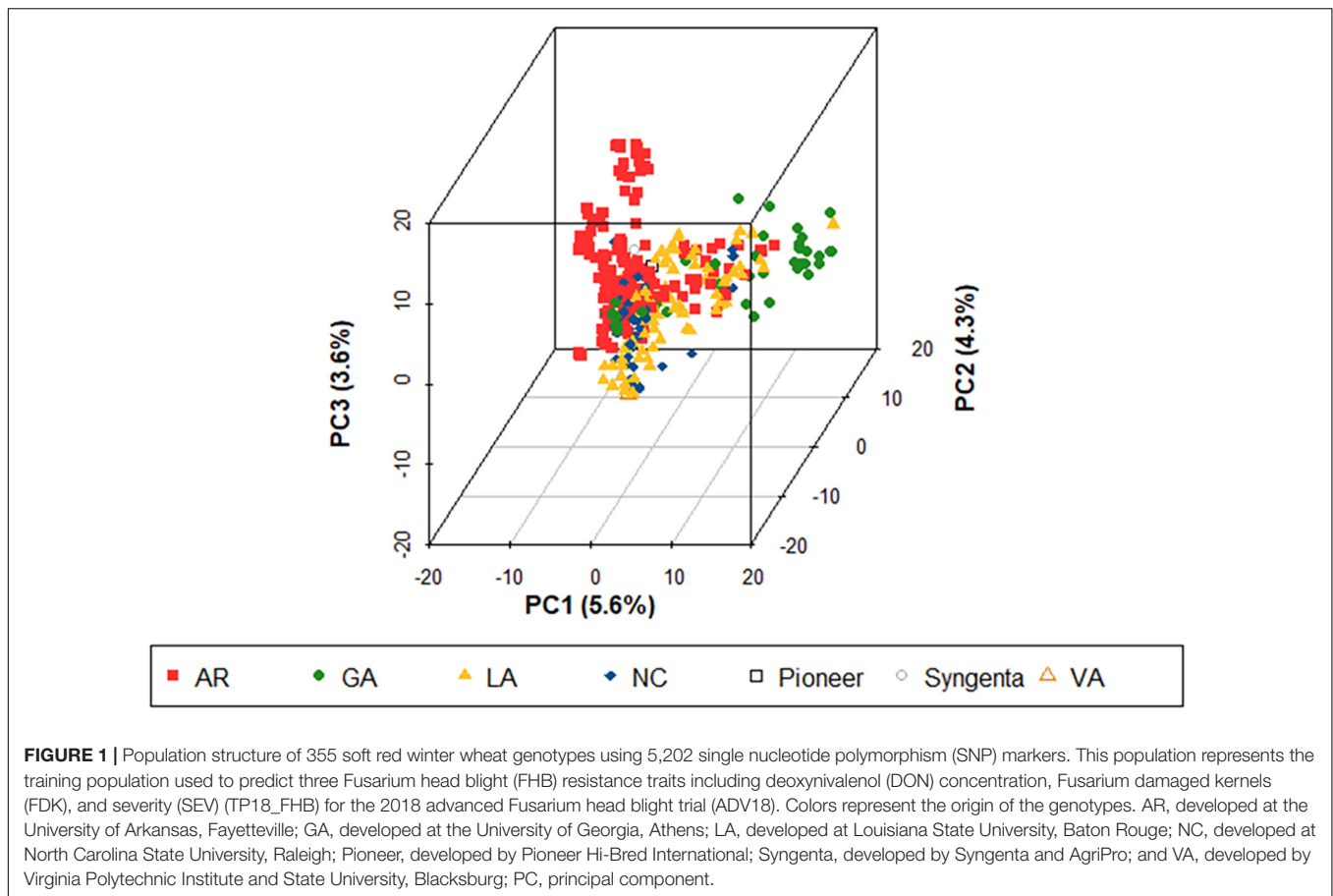
[†]ns, nonsignificant at the 0.05 probability level.

principal components only accounted for 5.23, 3.99, and 3.42% of the total genetic variation (**Figure 1**). There was no noticeable differentiation between the TP18_FHB population and ADV18 and the TP19_FHB population and ADV19 (**Supplementary Figures 1A,B**).

Cross Validation

Between both TPs, the MTGS models had significantly higher prediction accuracies compared to NGS models for DON, FDK, and SEV (**Figure 2**). Prediction accuracies for DON decreased between TP18_FHB and TP19_FHB while prediction

accuracies for FDK and SEV increased. The decrease in prediction accuracy for DON was likely a result of background population structure within TP19_FHB between genotypes from the TP18_FHB population, which does not contain genotypes with *Fhb1*, and ADV18 which does contain genotypes with *Fhb1* (**Supplementary Figure 1A**). The trait with the highest mean prediction accuracies among the NGS models for TP18_FHB was DON, with a mean accuracy of 0.61, while the trait with the highest prediction accuracy for TP19_FHB was SEV ($r = 0.61$). The trait with the second highest mean prediction accuracy among the NGS models for TP18_FHB was SEV ($r = 0.54$)



while DON and FDK had the same mean prediction accuracy for TP19_FHB ($r = 0.49$). Fusarium damaged kernels had the lowest mean prediction accuracy among the NGS models for TP18_FHB ($r = 0.45$). The ranking of traits between the MTGS models was not consistent with the NGS models or between TPs. Severity had the highest prediction accuracy in TP18_FHB ($r = 0.76$), followed by FDK ($r = 0.74$) and DON ($r = 0.72$). With TP19_FHB, DON also had the MTGS model with the lowest mean prediction accuracy ($r = 0.66$), while FDK and SEV had mean prediction accuracies of 0.74 (Figure 2).

Forward Prediction

When TP18_FHB was used to predict DON, FDK, and SEV for ADV18, there were significant correlations between the GEBVs calculated from the NGS and MTGS models and phenotypes for all FHB resistance traits. The strength of both correlations decreased for all methods when compared with phenotypic data from ARE19, with the exception for the MTGS model for SEV, where the correlation increased to $r = 0.60$ compared with $r = 0.57$ (Table 3). Both NGS and MTGS models had higher selection accuracies compared to phenotypic selection from ADV18 DON data (52.9%), where the NGS model correctly selected 82.4% of genotypes in ARE19, while the MTGS model correctly selected 70.6% (Table 3 and Figures 3A,B). The NGS ($R = -0.37 \mu\text{g g}^{-1}$) model had the highest response to selection for DON compared

to the NGS model ($R = -0.23 \mu\text{g g}^{-1}$) and phenotypic selection ($R = 0.20 \mu\text{g g}^{-1}$) (Table 3).

When predicting FDK for ADV18, the MTGS model had the strongest correlations with the ADV18 FDK data as well as the FDK adjusted means from ARE19. The NGS ($R = -4.09\%$) model again had the highest response to selection than the MTGS ($R = -2.83\%$) model and phenotypic selection ($R = -1.59\%$) for FDK (Table 3). The MTGS and NGS models had the same selection accuracy for FDK (70.6%) where both models outperformed phenotypic selection based on adjusted means for FDK from ADV18 (58.8%) (Table 3 and Figures 3C,D).

The MTGS model had stronger correlations between GEBVs for SEV and adjusted means for SEV from ADV18 and ARE19 than the NGS model (Table 3). The MTGS model also had the strongest response to selection ($R = -2.29\%$) and selection accuracy (47.1%) compared with the NGS model, where $R = -0.82\%$ and selection accuracy was 41.2%. The NGS model underperformed phenotypic selection for both response to selection ($R = -1.49\%$) and selection accuracy (52.9%), with the MTGS model only underperforming phenotypic selection for selection accuracy (Table 3 and Figures 3E,F).

When using TP19_FHB to predict FHB resistance traits for ADV19, the correlations between GEBVs from the MTGS models and phenotypic results from ADV19 were stronger than TP18_FHB for all three traits. Correlations between

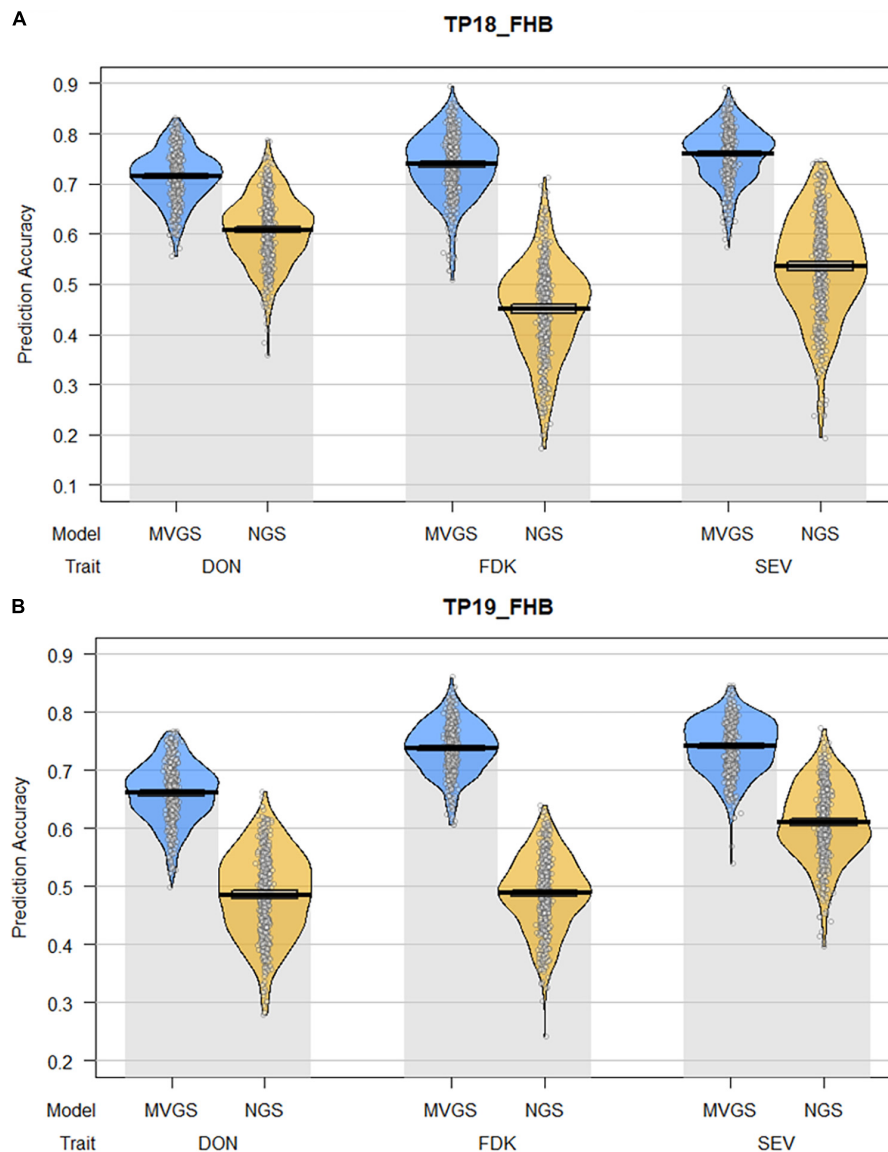


FIGURE 2 | Pirte plots comparing the mean prediction accuracies between multi-trait genomic selection (MTGS) models with naïve genomic selection (NGS) models for three Fusarium head blight resistance traits (FHB), deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) in soft red winter wheat across two training populations (TPs): **(A)** TP18_FHB, TP used to predict three FHB resistance traits for the 2018 advanced $F_{4:7}$ generation (ADV18); **(B)** TP19_FHB, TP used to predict three FHB resistance traits for the 2019 advanced $F_{4:7}$ generation (ADV19), consisting of all genotypes from TP18_FHB and ADV18. The x-axis represents the combination of FHB resistance traits and GS model used to predict each trait. The y-axis represents the mean prediction accuracy across 100 iterations of fivefold cross-validation in the form of a Pearson correlation coefficient (r) between the predicted genome-estimated breeding value (GEBV) and the actual phenotypic value for the validation populations. Individual points represent the Pearson correlation from each fold of each iteration of cross-validation for a total of 500 data points. The lines within each plot represent the mean and 95% confidence intervals for prediction accuracy. The curves represent the smoothed densities of the data.

GEBVs from the MTGS models were stronger than TP18_FHB when compared with adjusted means from ARE20 for DON and FDK (Table 3). Response to selection for TP19_FHB was different from TP18_FHB in that phenotypic selection outperformed the GS models for DON and SEV, whereas the MTGS model had a stronger response to selection than the NGS model and phenotypic selection for FDK (Table 3). Selection accuracies did change between TPs, as the MTGS model

(69.6%) outperformed both phenotypic selection (13.0%) and the NGS model (56.5%) for DON for TP19_FHB (Table 3 and Figures 4A,B). Unlike the results for TP18_FHB, both GS models had far lower selection accuracies than phenotypic selection (91.3%), although the MTGS model (60.9%) was better than the NGS model (34.8%) (Table 3 and Figures 4C,D). Selection accuracy for SEV also changed, where the MTGS model had the same selection accuracy as phenotypic selection (82.6%)

TABLE 3 | Comparison of three selection methods, phenotypic selection based on three FHB resistance traits using two training populations (TP), deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) from the advanced trials (ADV), naïve genomic selection (NGS), and multi-trait genomic selection (MTGS), based on correlations between genome estimated breeding values and the adjusted means from following generations, response to selection, and selection accuracy of genotypes in the final generation.

| TP | Trait | Method | r ADV ^a | r ARE ^b | Selection differential | Response to selection | Selection accuracy |
|----------|-------|--------|----------------------|----------------------|------------------------|-----------------------|--------------------|
| TP18_FHB | DON | PS | — | −0.01 ^{ns†} | 0.40 | 0.20 | 52.9 |
| | | NGS | 0.22* | 0.19 ^{ns} | −0.73 | −0.37 | 82.4 |
| | | MTGS | 0.53*** | 0.10 ^{ns} | −0.46 | −0.23 | 70.6 |
| | FDK | PS | — | 0.14 ^{ns} | −2.24 | −1.59 | 58.8 |
| | | NGS | 0.41*** | 0.38 ^{ns} | −5.77 | −4.09 | 70.6 |
| | | MTGS | 0.70*** | 0.42 ^{ns} | −3.99 | −2.83 | 70.6 |
| | SEV | PS | — | 0.54* | −3.46 | −1.49 | 52.9 |
| | | NGS | 0.29** | 0.16 ^{ns} | −1.90 | −0.82 | 41.2 |
| | | MTGS | 0.57*** | 0.60* | −5.33 | −2.29 | 47.1 |
| TP19_FHB | DON | PS | — | 0.51* | −1.32 | −1.03 | 13.0 |
| | | NGS | 0.17 ^{ns} | 0.37 ^{ns} | −0.67 | −0.53 | 56.5 |
| | | MTGS | 0.71*** | 0.45* | −0.96 | −0.75 | 69.6 |
| | FDK | PS | — | 0.67*** | −4.07 | −3.42 | 91.3 |
| | | NGS | 0.18* | 0.45* | −3.21 | −2.70 | 34.8 |
| | | MTGS | 0.83*** | 0.64** | −4.57 | −3.84 | 60.9 |
| | SEV | PS | — | 0.78*** | −5.86 | −4.45 | 82.6 |
| | | NGS | 0.25** | 0.08 ^{ns} | 0.50 | 0.38 | 60.9 |
| | | MTGS | 0.67*** | 0.12 ^{ns} | −0.18 | −0.13 | 82.6 |

^aPearson correlation coefficient between GEBVs and phenotypic data from the ADV population used as a validation population (VP).

^bPearson correlations coefficient between GEBVs and adjusted means for phenotypic data from the elite (ARE) generation.

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

[†]ns, nonsignificant at the 0.05 probability level.

while also outperforming the NGS model (60.9%) (Table 3 and Figures 4E,F).

DISCUSSION

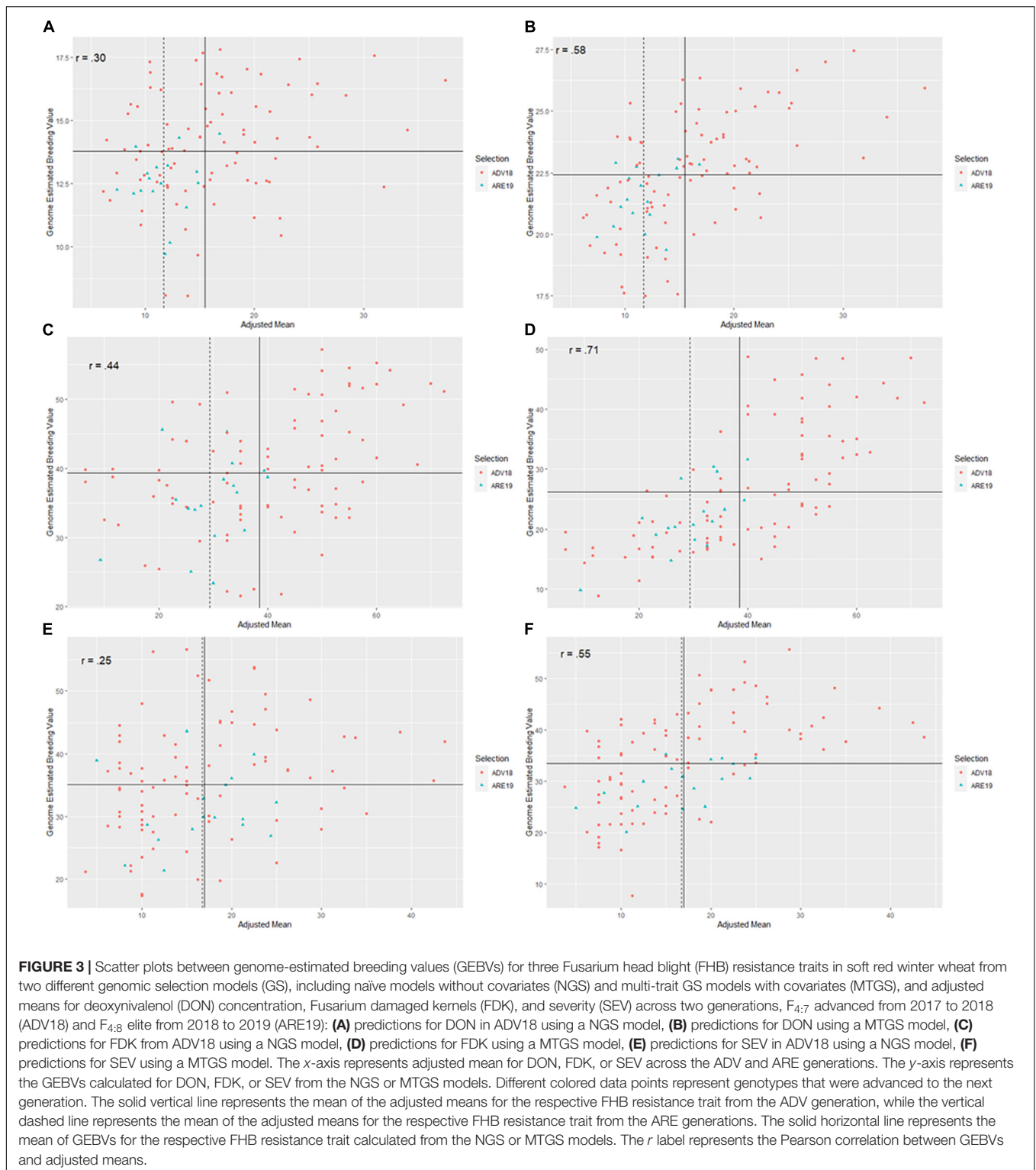
Genomic selection is a valuable tool for plant breeders, and many studies have shown the vast realm of possibilities for its application (Heffner et al., 2009; Sorrells, 2015; Larkin et al., 2019). The primary goal for GS is to increase genetic gain for a trait of interest within a breeding program through the reduction of time within a breeding cycle and by improving selection accuracy (Schaeffer, 2006; Bernardo and Yu, 2007; Heffner et al., 2009; Asoro et al., 2013; Rutkoski et al., 2015). While most research in GS has focused on optimizing TPs to increase model predictive ability, less have focused on the implementation of GS into breeding programs in the form of forward selection (Bernardo, 2016). In our study, we chose to focus on forward prediction using NGS and MTGS models and compared their performance, based on selection accuracy and response to selection, to phenotypic selection for economically important traits, such as FHB resistance.

Prediction Accuracy of Training Populations

In our study, we saw that MTGS models consistently had significantly higher prediction accuracies for DON, FDK, and

SEV in every TP compared to NGS. These results were consistent with previous studies involving MTGS for FHB resistance traits (Schulthess et al., 2018; Larkin et al., 2020; Moreno-Amores et al., 2020). This follows the general trend for MTGS, where covariate traits sharing a strong correlation with a trait of interest can improve prediction accuracies for said trait of interest (Calus and Veerkamp, 2011; Jia and Jannink, 2012; Schulthess et al., 2016; Lozada and Carter, 2019; Ward et al., 2019).

Regarding the correlations between FHB resistance traits, it is interesting to note that HD was consistently negatively correlated with SEV, and yet positively correlated with DON. The negative correlation between SEV and HD has been observed in many different studies (Gervais et al., 2003; Paillard et al., 2004; Schmolke et al., 2005; Larkin et al., 2020; Moreno-Amores et al., 2020). This is because wheat genotypes that flower earlier are exposed to more favorable conditions for FHB infection, such as higher humidity and rainfall during the early growing season, versus the later part of the growing season (Buerstmayr et al., 2019). However, while positive correlations have been observed between HD and DON in other studies, less is known about this association (Liu et al., 2012; Agnes et al., 2014; Larkin et al., 2020). Agnes et al. (2014) suggested that this positive correlation was related to additional fungal growth after the soft dough stage (Feekes 11.2). Several groups have also identified QTL associated with both DON and HD (Schmolke et al., 2005; Lin et al., 2008; Agnes et al., 2014). Agnes et al. (2014) specifically identified such a QTL on chromosome 7B, which was co-located



with the vernalization response gene *Vrn-B3*. Even so, like most FHB resistance traits and HD, we believe that this association is variable and environmentally dependent (Buerstmayr et al., 2019), as we saw correlations between DON and HD ranging between $r = 0.01$ and $r = 0.31$ (Table 2).

We also updated our TPs for each generation by adding phenotypic data for genotypes from the previous generation into the following year's TP. Other studies have found that updating TPs helped to prevent the deviation in genetic relationships between the TP and VP as new germplasm was added and

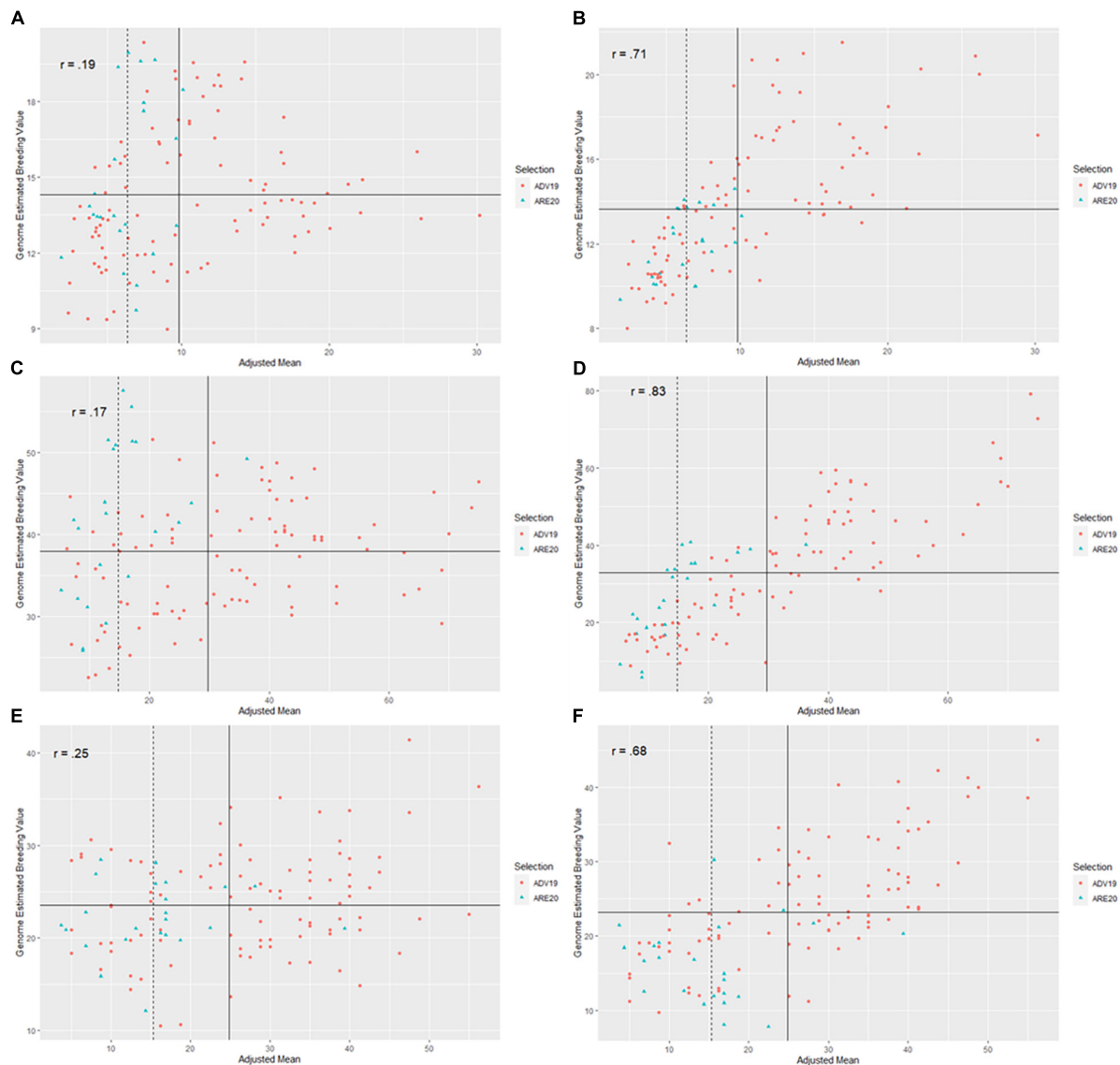


FIGURE 4 | Scatter plots between genome-estimated breeding values (GEBVs) for three Fusarium head blight (FHB) resistance traits in soft red winter wheat from two different genomic selection models (GS), including naïve models without covariates (NGS) and multi-trait GS models with covariates (MTGS), and adjusted means for deoxynivalenol (DON) concentration, Fusarium damaged kernels (FDK), and severity (SEV) across two generations, F_{4:7} advanced from 2018 to 2019 (ADV19) and F_{4:8} elite from 2019 to 2020 (ARE20): **(A)** predictions for DON in ADV19 using a NGS model, **(B)** predictions for DON using a MTGS model, **(C)** predictions for FDK from ADV19 using a NGS model, **(D)** predictions for FDK using a MTGS model, **(E)** predictions for SEV in ADV18 using a NGS model, **(F)** predictions for SEV using a MTGS model. The x-axis represents adjusted mean for DON, FDK, or SEV across the ADV and ARE generations. The y-axis represents the GEBVs calculated for DON, FDK, or SEV from the NGS or MTGS models. Different colored data points represent genotypes that were advanced to the next generation. The solid vertical line represents the mean of the adjusted means for the respective FHB resistance trait from the ADV generation, while the vertical dashed line represents the mean of the adjusted means for the respective FHB resistance trait from the ARE generations. The solid horizontal line represents the mean of GEBVs for the respective FHB resistance trait calculated from the NGS or MTGS models. The r label represents the Pearson correlation between GEBVs and adjusted means.

advanced through the breeding program (Meuwissen, 2009; Clark et al., 2012; Lorenz et al., 2012; Lorenz and Smith, 2015; Neyhart et al., 2017). Studies have also shown that larger TP sizes can have higher prediction accuracies as well, particularly when working with more diverse populations where new germplasm is continually added to the breeding program (Heffner et al., 2011; Mujibi et al., 2011; Heslot et al., 2012; Poland et al., 2012;

Isidro et al., 2015; Norman et al., 2018). We also observed this trend for FDK and SEV between TP18_FHB and TP19_FHB; however, we did not observe this trend for DON, where prediction accuracy decreased when additional genotypes were added from ADV18. This can likely be attributed to less variation and a lower heritability for DON within ADV18. Genotypes within ADV18 also had the FHB resistance alleles for *Fhb1*,

which could have increased background population structure within TP19_FHB.

Forward Prediction

Much like the results from the cross-validation analyses of the TPs, the MTGS models had stronger correlations between their calculated GEBVs and phenotypic results from their respective VPs for FHB resistance traits, aligning with other studies involving MTGS models (Jia and Jannink, 2012; Schulthess et al., 2016; Lozada and Carter, 2019; Ward et al., 2019; Larkin et al., 2020). This was clearly observed with TP18_FHB, when correlations between MTGS GEBVs and ADV18 phenotypic results were compared with correlations between NGS GEBVs and ADV18 phenotypic results for all three traits. The prediction accuracy advantage of the MTGS model was also observed with correlations between GEBVs and ARE19 phenotypic results for FDK and SEV when compared with NGS.

Our range in prediction accuracy for the NGS models were between $r = 0.08$ and $r = 0.45$ while the range of our MTGS models was between $r = 0.10$ and $r = 0.83$. These prediction accuracies were within the range of prediction accuracies observed for FHB resistance traits in previous studies (Rutkoski et al., 2012; Arruda et al., 2015, 2016a; Larkin et al., 2020). However, the observation of lower prediction accuracies under specific circumstances was consistent with other studies with forward prediction for GY (Belamkar et al., 2018; Calvert et al., 2020). In an evaluation of forward prediction in the Kansas State University wheat breeding program, the highest prediction accuracy between the GEBVs for GY in the preliminary yield trials (PYTs) and the actual phenotypic results for GY was $r = -0.16$ (Calvert et al., 2020). The same study also used high-throughput phenotyping traits as covariates in a MTGS model for forward prediction of GY in wheat, however, the prediction accuracy was unfavorable unless a large TP was used (Calvert et al., 2020). This contrasts with our results where the use of other FHB resistance or agronomic traits as covariates significantly improved prediction accuracy for both TPs.

The MTGS model was also superior to phenotypic selection based on ADV18 phenotypic data for all three traits; however, this advantage disappeared when implementing the models trained with TP19_FHB. This is likely because genotypes in ADV19 had a much higher prevalence of resistance alleles for *Fhb1* compared with TP19_FHB, therefore the TP failed to account for this major source of genetic resistance to FHB in the VP. This highlights the importance of the TP being able to account for population structure existing within the VP, otherwise prediction accuracies can be lower. Such a result was foreshadowed with the lower prediction accuracies from the cross-validation for TP19_FHB, where no genotypes from the initial TP18_FHB contained resistance alleles for *Fhb1*, while only a small portion of genotypes from ADV18 contained the resistance alleles. A more detailed description of major and minor FHB resistance QTL present within TP18_FHB can be found in Larkin et al. (2020).

Response to selection was measured as the difference between the mean of the top 50% of breeding lines in the ARE generation, selected based on GEBVs and adjusted means of FHB resistance traits for the ADV population, compared with the mean of the

full ARE population. Other studies have shown that GS could not have as high of a response to selection as phenotypic selection; however, our method of excluding phenotypic data from the ADV genotypes from the selection dataset allowed for greater independence from bias toward the phenotypic selection method (Lozada et al., 2019). In terms of response to selection, both GS models were superior to phenotypic selection for DON and FDK, and the MTGS model for SEV, when using the TP18_FHB to predict ADV18. Much like the results for prediction accuracy, this strong advantage was not observed when using the TP19_FHB to predict ADV19, except for the MTGS model for FDK, likely due to the same reasons described above. There have been no extensive forward prediction studies for FHB resistance traits in wheat. Regardless, the fact that phenotypic selection did not significantly outperform the MTGS model across years or traits indicates that MTGS models may be a good supplement, if not substitute for phenotypic selection, particularly during years when it is difficult to phenotype.

When comparing GS models with phenotypic selection for FHB resistance traits based on selection accuracy, the NGS and MTGS models had higher selection accuracies for DON using TP18_FHB, and the MTGS model was equal to phenotypic selection using TP19_FHB. Both the MTGS and NGS models were equally more accurate than phenotypic selection for FDK with TP18_FHB. Additionally, the MTGS model was equal in performance with phenotypic selection for SEV in TP19_FHB. It has been mentioned that prediction accuracy does not necessarily correlate with selection accuracy for forward prediction (Belamkar et al., 2018).

CONCLUSION

This study showed that both NGS and MTGS could be successfully implemented into a SRWW breeding program, while using other agronomic and disease traits as covariates with reasonable accuracy compared to phenotypic selection and again asserted its value as a tool for plant breeders. We also found that MTGS models performed significantly better than NGS models in terms of both cross-validation within TPs as well as forward prediction of untested genotypes for economically important traits, such as FHB resistance traits. This was particularly evident when there was a strong correlation between the trait of interest and the covariate trait. This is one of the first studies to show that MTGS could be effectively implemented for forward prediction within a wheat breeding program. This is also the first study to extensively investigate the use of forward prediction when breeding for FHB resistance in wheat. We found that GS could serve as a suitable, albeit imperfect, alternative to phenotypic selection when implemented during years where environmental conditions prohibit accurate phenotypic selection, particularly when experiencing late freezing events or extensive lodging.

Prior to implementing GS into their own breeding programs, breeders must consider the genetic relationships between their prospective TPs and the breeding lines they hope to use as their VP. In the case of MTGS, breeders must also consider the correlations between their traits of interest and secondary traits

used as covariates, as these correlations can differ between the TP and VP. For example, there could be a strong correlation between DON and HD in the TP but there could be a weak correlation between the two traits in the VP, therefore the MTGS model might not be more accurate than a NGS model. Inversely, there could be a strong correlation between traits in the VP while there is a weak correlation between traits in the TP, therefore MTGS could be more accurate than expected when forward prediction is implemented.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the FigShare repository, accession numbers <https://doi.org/10.6084/m9.figshare.16685902.v1>, <https://doi.org/10.6084/m9.figshare.16685701.v2>, <https://doi.org/10.6084/m9.figshare.16685797.v1>, <https://doi.org/10.6084/m9.figshare.16685722.v1>, and <https://doi.org/10.6084/m9.figshare.16685707.v1>.

AUTHOR CONTRIBUTIONS

DL and RM conceived and designed the experiments. DL performed data analyses, conducted the experiments, and

wrote the manuscript. DM and AH collected data from the disease nurseries. RM, BW, AH, DM, and GB-G edited the manuscript. GB-G and BW conducted genotyping and generated hapmaps. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the USDA-ARS, under Agreement no. USDA-ARS 59-0206-7-005, a cooperative project with the U.S. Wheat and Barley Scab Initiative (USWBSI), the Agriculture and Food Research Initiative (AFRI) of the USDA National Institute of Food and Agriculture (NIFA) Grant 2017-67007-25939 (Wheat-CAP) and in collaboration with SunGrains. Individual state commodity boards also supported this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.715314/full#supplementary-material>

REFERENCES

- Agnes, S. H., Szabolcs, L. K., Monika, V., Laszlo, P., Janos, P., Csaba, L., et al. (2014). Differential influence of QTL linked to *Fusarium* head blight, *Fusarium*-damaged kernel, deoxynivalenol contents and associated morphological traits in a Frontana-derived wheat population. *Euphytica* 200, 9–26. doi: 10.1007/s10681-014-1124-2
- Akdemir, D., Sanchez, J. I., and Jannink, J. L. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genet. Sel. Evol.* 47:38. doi: 10.1186/s12711-015-0116-6
- 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. doi: 10.1126/science.aar7191
- Argyris, J., Van Sanford, D., and TeKrony, D. (2003). *Fusarium graminearum* infection during wheat seed development and its effect on seed quality. *Crop Sci.* 43, 1782–1788. doi: 10.2135/cropsci2003.1782
- Arruda, M. P., Brown, P. J., Lipka, A. E., Krill, A. M., Thurber, C., and Kolb, F. L. (2015). Genomic selection for predicting *Fusarium* head blight resistance in a wheat breeding program. *Plant Genome* 8:12. doi: 10.3835/plantgenome2015.01.0003
- Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., et al. (2016a). Genome-wide association mapping of *Fusarium* head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome* 9, 1–14. doi: 10.3835/plantgenome2015.04.0028
- Arruda, M. P., Lipka, A. E., Brown, P. J., Krill, A. M., Thurber, C., Brown-Guedira, G., et al. (2016b). Comparing genomic selection and marker-assisted selection for *Fusarium* head blight resistance in wheat (*Triticum aestivum* L.). *Mol. Breed.* 36:84. doi: 10.1007/s11032-016-0508-5
- Asoro, F. G., Newell, M. A., Beavis, W. D., Scott, M. P., Tinker, N. A., and Jannink, J. L. (2013). Genomic, marker-assisted, and pedigree-BLUP selection methods for beta-glucan concentration in elite oat. *Crop Sci.* 53, 1894–1906. doi: 10.2135/cropsci2012.09.0526
- Belamkar, V., Guttieri, M. J., Hussain, W., Jarquin, D., El-basyoni, I., Poland, J., et al. (2018). Genomic selection in preliminary yield trials in a winter wheat breeding program. *G3* 8, 2735–2747. doi: 10.1534/g3.118.200415
- Benson, J., Brown-Guedira, G., Murphy, J. P., and Sneller, C. (2012). Population structure, linkage disequilibrium, and genetic diversity in soft winter wheat enriched for *Fusarium* head blight resistance. *Plant Genome* 5, 71–80. doi: 10.3835/plantgenome2011.11.0027
- Bernardo, R. (2016). Bandwagons I, too, have known. *Theor. Appl. Genet.* 129, 2323–2332. doi: 10.1007/s00122-016-2772-5
- Bernardo, R., and Yu, J. M. (2007). Prospects for genomewide selection for quantitative traits in maize. *Crop Sci.* 47, 1082–1090. doi: 10.2135/cropsci2006.11.0690
- Boyles, R. E., Marshall, D. S., and Bockelman, H. E. (2019). Yield data from the uniform southern soft red winter wheat nursery emphasize importance of selection location and environment for cultivar development. *Crop Sci.* 59, 1887–1898. doi: 10.2135/cropsci2018.11.0685
- 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
- Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi: 10.1016/j.ajhg.2018.07.015
- Buerstmayr, M., Steiner, B., and Buerstmayr, H. (2019). Breeding for *Fusarium* head blight resistance in wheat-progress and challenges. *Plant Breed.* 139, 429–454. doi: 10.1111/pbr.12797
- Calus, M. P. L., and Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genet. Sel. Evol.* 43:26. doi: 10.1186/1297-9686-43-26
- Calvert, M., Evers, B., Wang, X., Fritz, A., and Poland, J. (2020). Breeding program optimization for genomic selection in winter wheat. *bioRxiv* [Preprint] doi: 10.1101/2020.10.07.330415
- Clark, S. A., Hickey, J. M., Daetwyler, H. D., and van der Werf, J. H. J. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genet. Sel. Evol.* 44:4. doi: 10.1186/1297-9686-44-4
- Combs, E., and Bernardo, R. (2013). Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. *Plant Genome* 6:7. doi: 10.3835/plantgenome2012.11.0030
- Covarrubias-Pazarán, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., et al. (2018). Multivariate GBLUP improves accuracy of genomic selection for yield and fruit weight in biparental populations of

- Vaccinium macrocarpon* ait. *Front. Plant Sci.* 9:1310. doi: 10.3389/fpls.2018.01310
- Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., and Poland, J. (2018). Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding. *Plant Genome* 11:170043. doi: 10.3835/plantgenome2017.05.0043
- Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024
- Endelman, J. B., and Jannink, J. L. (2012). Shrinkage estimation of the realized relationship matrix. *G3* 2, 1405–1413. doi: 10.1534/g3.112.004259
- Falconer, D. S., and McKay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th Edn. Harlow: Longmans Green.
- FDA (2010). *Guidance for Industry and FDA: Advisory Levels for Deoxynivalenol (DON) in Finished Wheat Products for Human Consumption and Grains and Grain By-Products used for Animal Feed*. Silver Spring, MD: FDA.
- Gervais, L., Dedryver, F., Morlais, J. Y., Bodusseau, V., Negre, S., Bilous, M., et al. (2003). Mapping of quantitative trait loci for field resistance to *Fusarium* head blight in an European winter wheat. *Theor. Appl. Genet.* 106, 961–970. doi: 10.1007/s00122-002-1160-5
- Goral, T., Wisniewska, H., Ochodzki, P., Nielsen, L. K., Walentyn-Goral, D., and Stepień, L. (2019). Relationship between *Fusarium* head blight, kernel damage, concentration of *Fusarium* biomass, and *Fusarium* toxins in grain of winter wheat inoculated with *Fusarium culmorum*. *Toxins* 11:2. doi: 10.3390/toxins11010002
- Guo, G., Zhao, F. P., Wang, Y. C., Zhang, Y., Du, L. X., and Su, G. S. (2014). Comparison of single-trait and multiple-trait genomic prediction models. *BMC Genet.* 15:30. doi: 10.1186/1471-2156-15-30
- Guo, J., Khan, J., Pradhan, S., Shahi, D., Khan, N., Avci, M., et al. (2020). Multi-trait genomic prediction of yield-related traits in US soft wheat under variable water regimes. *Genes* 11:1270. doi: 10.3390/genes11111270
- Habier, D., Fernando, R. L., and Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics* 177, 2389–2397. doi: 10.1534/genetics.107.081190
- Harrison, S. A., Babar, M. A., Barnett, R. D., Blount, A. R., Johnson, J. W., Mergoum, M., et al. (2017). 'LA05006', a dual-purpose oat for louisiana and other southeastern regions of the USA. *J. Plant Regist.* 11, 89–94. doi: 10.3198/jpr2016.08.0040crc
- Heffner, E. L., Jannink, J. L., Iwata, H., Souza, E., and Sorrells, M. E. (2011). Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* 51, 2597–2606. doi: 10.2135/cropsci2011.05.0253
- 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
- Heslot, N., Yang, H. P., Sorrells, M. E., and Jannink, J. L. (2012). Genomic selection in plant breeding: a comparison of models. *Crop Sci.* 52, 146–160. doi: 10.2135/cropsci2011.06.0297
- Isidro, J., Jannink, J. L., Akdemir, D., Poland, J., Heslot, N., and Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *Theor. Appl. Genet.* 128, 145–158. doi: 10.1007/s00122-014-2418-4
- Jannink, J. L., Lorenz, A. J., and Iwata, H. (2010). Genomic selection in plant breeding: from theory to practice. *Brief. Funct. Genomics* 9, 166–177. doi: 10.1093/bfpg/eq001
- Jia, H. Y., Zhou, J. Y., Xue, S. L., Li, G. Q., Yan, H. S., Ran, C. F., et al. (2018). A journey to understand wheat *Fusarium* head blight resistance in the Chinese wheat landrace Wangshuibai. *Crop J.* 6, 48–59. doi: 10.1016/j.cj.2017.09.006
- Jia, Y., and Jannink, J. L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* 192, 1513–1522. doi: 10.1534/genetics.112.144246
- Johnson, J. W., Chen, Z., Buck, J. W., Buntin, G. D., Babar, M. A., Mason, R. E., et al. (2017). 'GA 03564-12E6': a high-yielding soft red winter wheat cultivar adapted to georgia and the southeastern regions of the United States. *J. Plant Regist.* 11, 159–164. doi: 10.3198/jpr2016.07.0036crc
- Kruijer, W., Boer, M. P., Malosetti, M., Flood, P. J., Engel, B., Kooke, R., et al. (2015). Marker-based estimation of heritability in immortal populations. *Genetics* 199, 379–393. doi: 10.1534/genetics.114.167916
- Larkin, D. L., Holder, A. L., Mason, R. E., Moon, D. E., Brown-Guedira, G., Price, P. P., et al. (2020). Genome-wide analysis and prediction of *Fusarium* head blight resistance in soft red winter wheat. *Crop Sci.* 60, 2882–2900. doi: 10.1002/csc2.20273
- Larkin, D. L., Lozada, D. N., and Mason, R. E. (2019). Genomic selection-considerations for successful implementation in wheat breeding programs. *Agronomy* 9:479. doi: 10.3390/agronomy9090479
- 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
- Ligges, U., Maechler, M., Schnackenberg, S., and Ligges, M. U. (2018). *Package 'scatterplot3d'. Recuperado de https://cran.rproject.org/web/packages/scatterplot3d/scatterplot3d.pdf.*
- Lin, F., Xue, S. L., Tian, D. G., Li, C. J., Cao, Y., Zhang, Z. Z., et al. (2008). Mapping chromosomal regions affecting flowering time in a spring wheat RIL population. *Euphytica* 164, 769–777. doi: 10.1007/s10681-008-9724-3
- Liu, S. Y., Christopher, M. D., Griffey, C. A., Hall, M. D., Gundrum, P. G., and Brooks, W. S. (2012). Molecular characterization of resistance to *Fusarium* head blight in U.S. soft red winter wheat breeding line VA00W-38. *Crop Sci.* 52, 2283–2292. doi: 10.2135/cropsci2012.03.0144
- Lorenz, A. J., and Smith, K. P. (2015). Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. *Crop Sci.* 55, 2657–2667. doi: 10.2135/cropsci2014.12.0827
- Lorenz, A. J., Smith, K. P., and Jannink, J. L. (2012). Potential and optimization of genomic selection for *Fusarium* head blight resistance in six-row barley. *Crop Sci.* 52, 1609–1621. doi: 10.2135/cropsci2011.09.0503
- Lozada, D. N., and Carter, A. H. (2019). Accuracy of single and multi-trait genomic prediction models for grain yield in US Pacific northwest winter wheat. *Crop Breed. Genet. Genomics* 1:e190012. doi: 10.20900/cbgg20190012
- Lozada, D. N., Mason, R. E., Sarinelli, J. M., and Brown-Guedira, G. (2019). Accuracy of genomic selection for grain yield and agronomic traits in soft red winter wheat. *BMC Genet.* 20:82. doi: 10.1186/s12863-019-0785-1
- Lozada, D. N., Ward, B. P., and Carter, A. H. (2020). Gains through selection for grain yield in a winter wheat breeding program. *PLoS One* 15:e0221603. doi: 10.1371/journal.pone.0221603
- Mason, R. E., Johnson, J. W., Mergoum, M., Miller, R. G., Moon, D. E., Carlin, J. F., et al. (2018). AR11LE24, a soft red winter wheat adapted to the mid-south region of the USA. *J. Plant Regist.* 12, 357–361. doi: 10.3198/jpr2017.09.0060crc
- Massman, J. M., Jung, H. J. G., and Bernardo, R. (2013). Genomewide selection versus marker-assisted recurrent selection to improve grain yield and stover-quality traits for cellulosic ethanol in maize. *Crop Sci.* 53, 58–66. doi: 10.2135/cropsci2012.02.0112
- Mesterhazy, A. (1995). Types and components of resistance to *Fusarium* head blight of wheat. *Plant Breed.* 114, 377–386. doi: 10.1111/j.1439-0523.1995.tb00816.x
- Meuwissen, T. H. E. (2009). Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genet. Sel. Evol.* 41:35. doi: 10.1186/1297-9686-41-35
- Meuwissen, T. H. E., 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
- Michel, S., Ametz, C., Gungor, H., Akgol, B., Epure, D., Grausgruber, H., et al. (2017). Genomic assisted selection for enhancing line breeding: merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theor. Appl. Genet.* 130, 363–376. doi: 10.1007/s00122-016-2818-8
- Michel, S., Ametz, C., Gungor, H., Epure, D., Grausgruber, H., Loschenberger, F., et al. (2016). Genomic selection across multiple breeding cycles in applied bread wheat breeding. *Theor. Appl. Genet.* 129, 1179–1189. doi: 10.1007/s00122-016-2694-2
- Moreno-Amores, J., Michel, S., Miedaner, T., Longin, C. F. H., and Buerstmayr, H. (2020). Genomic predictions for *Fusarium* head blight resistance in a diverse durum wheat panel: an effective incorporation of plant height and heading date as covariates. *Euphytica* 216:19. doi: 10.1007/s10681-019-2551-x
- Mujibi, F. D. N., Nkrumah, J. D., Durunna, O. N., Stothard, P., Mah, J., Wang, Z., et al. (2011). Accuracy of genomic breeding values for residual feed intake in crossbred beef cattle. *J. Anim. Sci.* 89, 3353–3361. doi: 10.2527/jas.2010-3361
- Neyhart, J. L., Tiede, T., Lorenz, A. J., and Smith, K. P. (2017). Evaluating methods of updating training data in long-term genomewide selection. *G3* 7, 1499–1510. doi: 10.1534/g3.117.040550

- Norman, A., Taylor, J., Edwards, J., and Kuchel, H. (2018). Optimising genomic selection in wheat: effect of marker density, population size and population structure on prediction accuracy. *G3* 8, 2889–2899. doi: 10.1534/g3.118.200311
- Nyquist, W. (1962). Differential fertilization in the inheritance of stem rust resistance in hybrids involving a common wheat strain derived from *Triticum timopheevi*. *Genetics* 47:1109. doi: 10.1093/genetics/47.8.1109
- Paillard, S., Schnurbusch, T., Tiwari, R., Messmer, M., Winzeler, M., Keller, B., et al. (2004). QTL analysis of resistance to *Fusarium* head blight in Swiss winter wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 109, 323–332. doi: 10.1007/s00122-004-1628-6
- Phillips, N. D. (2017). Yarr! The pirate's guide to R. *APS Obs.* 30, 151–163.
- Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S. Y., Manes, Y., et al. (2012). Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5, 103–113. doi: 10.3835/plantgenome2012.06.0006
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rutkoski, J., Benson, J., Jia, Y., Brown-Guedira, G., Jannink, J. L., and Sorrells, M. (2012). Evaluation of genomic prediction methods for *Fusarium* head blight resistance in wheat. *Plant Genome* 5, 51–61. doi: 10.3835/plantgenome2012.02.0001
- Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Perez, L. G., Crossa, J., et al. (2016). Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3* 6, 2799–2808. doi: 10.1534/g3.116.032888
- Rutkoski, J., Singh, R. P., Huerta-Espino, J., Bhavani, S., Poland, J., Jannink, J. L., et al. (2015). Genetic gain from phenotypic and genomic selection for quantitative resistance to stem rust of wheat. *Plant Genome* 8:10. doi: 10.3835/plantgenome2014.10.0074
- Sarinelli, J. M., Murphy, J. P., Tyagi, P., Holland, J. B., Johnson, J. W., Mergoum, M., et al. (2019). Training population selection and use of fixed effects to optimize genomic predictions in a historical USA winter wheat panel. *Theor. Appl. Genet.* 132, 1247–1261. doi: 10.1007/s00122-019-03276-6
- Schaeffer, L. R. (2006). Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123, 218–223. doi: 10.1111/j.1439-0388.2006.00595.x
- Schmolke, M., Zimmermann, G., Buerstmayr, H., Schweizer, G., Miedaner, T., Korzun, V., et al. (2005). Molecular mapping of *Fusarium* head blight resistance in the winter wheat population Dream/Lynx. *Theor. Appl. Genet.* 111, 747–756. doi: 10.1007/s00122-005-2060-2
- Schroeder, H. W., and Christensen, J. J. (1963). Factors affecting resistance Of wheat to scab caused by *Gibberella Zeae*. *Phytopathology* 53, 831–838.
- Schulthess, A. W., Wang, Y., Miedaner, T., Wilde, P., Reif, J. C., and Zhao, Y. S. (2016). Multiple-trait- and selection indices-genomic predictions for grain yield and protein content in rye for feeding purposes. *Theor. Appl. Genet.* 129, 273–287. doi: 10.1007/s00122-015-2626-6
- Schulthess, A. W., Zhao, Y. S., Longin, C. F. H., and Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for *Fusarium* head blight severity in hybrid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 131, 685–701. doi: 10.1007/s00122-017-3029-7
- Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., and Kizek, R. (2010). Deoxynivalenol and its toxicity. *Interdiscip. Toxicol.* 3, 94–99. doi: 10.2478/v10102-010-0019-x
- Sorrells, M. E. (2015). “Genomic selection in plants: empirical results and implications for wheat breeding,” in *Advances in Wheat Genetics: From Genome to Field*, eds Y. Ogihara, S. Takumi, and H. Handa (Tokyo: Springer), 401–409. doi: 10.1007/978-4-431-55675-6_45
- Steiner, B., Michel, S., Maccaferri, M., Lemmens, M., Tuberosa, R., and Buerstmayr, H. (2019). Exploring and exploiting the genetic variation of *Fusarium* head blight resistance for genomic-assisted breeding in the elite durum wheat gene pool. *Theor. Appl. Genet.* 132, 969–988. doi: 10.1007/s00122-018-3253-9
- Sun, J., Rutkoski, J. E., Poland, J. A., Crossa, J., Jannink, J. L., and Sorrells, M. E. (2017). Multitrait, random regression, or simple repeatability model in high-throughput phenotyping data improve genomic prediction for wheat grain yield. *Plant Genome* 10:12. doi: 10.3835/plantgenome2016.11.0111
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980
- Ward, B. P., Brown-Guedira, G., Tyagi, P., Kolb, F. L., Van Sanford, D. A., Sneller, C. H., et al. (2019). Multienvironment and multitrait genomic selection models in unbalanced early-generation wheat yield trials. *Crop Sci.* 59, 491–507. doi: 10.2135/cropsci2018.03.0189
- Wickham, H., Chang, W., and Wickham, M. H. (2016). Package ‘ggplot2’. Create elegant data visualisations using the grammar of graphics. *Version 2*, 1–189.
- Wilson, W., McKee, G., Nganje, W., Dahl, B., and Bangsund, D. (2017). *Economic Impact of USWBSI's Scab Initiative to Reduce FHB*. Fargo, ND: North Dakota State University.
- Xu, Y. B., and Crouch, J. H. (2008). Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci.* 48, 391–407. doi: 10.2135/cropsci2007.04.0191

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 Larkin, Mason, Moon, Holder, Ward and Brown-Guedira. 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.



Investigation and Genome-Wide Association Analysis of Fusarium Seedling Blight Resistance in Chinese Elite Wheat Lines

Yike Liu^{1,2}, Guang Zhu¹, Zhangwang Zhu¹, Lin Chen¹, Hongli Niu^{3,4}, Weijie He¹, Hanwen Tong¹, Jinghan Song¹, Yuqing Zhang¹, Dongfang Ma^{2,3,4*} and Chunbao Gao^{1,2*}

¹ Hubei Key Laboratory of Food Crop Germplasm and Genetic Improvement, Food Crops Institute, Hubei Engineering and Technology Research, Hubei Academy of Agricultural Sciences, Wuhan, China, ² Center of Wheat, Wheat Disease Biology Research Station for Central China, Wuhan, China, ³ Engineering Research Center of Ecology and Agricultural Use of Wetland, Ministry of Education, Yangtze University, Jingzhou, China, ⁴ Hubei Collaborative Innovation Center for Grain Industry, College of Agriculture, Yangtze University, Jingzhou, China

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Mohsen Mohammadi,
Purdue University, United States
Shengjie Liu,
Northwest A&F University, China

*Correspondence:

Dongfang Ma
madongfang1984@163.com
Chunbao Gao
gcbgybjw@163.com

Specialty section:

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

Received: 15 September 2021

Accepted: 18 October 2021

Published: 17 November 2021

Citation:

Liu Y, Zhu G, Zhu Z, Chen L,
Niu H, He W, Tong H, Song J,
Zhang Y, Ma D and Gao C (2021)
Investigation and Genome-Wide
Association Analysis of Fusarium
Seedling Blight Resistance in Chinese
Elite Wheat Lines.
Front. Plant Sci. 12:777494.
doi: 10.3389/fpls.2021.777494

Fusarium seedling blight (FSB) is an important disease of wheat occurring as part of the Fusarium disease complex consisting also of Fusarium head blight (FHB). 240 Chinese elite cultivars and lines were evaluated in greenhouse experiments for FSB resistance and genotyped using the wheat 90 K single nucleotide polymorphism arrays. Among them, 23 accessions had an average lesion length of less than 0.6 cm, exhibiting potential for breeding for FSB resistance in wheat. Jingfuma 1 and Yangmai 11 had a relatively high resistance to both FSB and FHB simultaneously. Six relatively stable quantitative trait loci (QTLs) were detected on chromosome arms 1DL, 3AS, 3BL, 6BL, 7AL, and Un using the mixed linear model approach, interpreting 4.83–7.53% of phenotypic variation. There was a negative correlation between the average FSB lesion length and the BLUE FHB index with a low coefficient, and resistance to both diseases appeared to be conferred by different QTLs across the same population. Four KASP markers were detected on 1DL, 3AS, 3BL, and 6BL in QTLs to facilitate marker-assisted selection. Combined with transcriptome data analysis, eight defense-related genes were considered as candidates for mapping QTLs. The resistant elite germplasm, mapped QTLs, and KASP markers developed in this study are useful resources for enhancing Fusarium seedling blight in wheat breeding.

Keywords: common wheat, Fusarium seedling blight, Fusarium head blight, GWAS, QTL

INTRODUCTION

Fusarium seedling blight (FSB) and Fusarium head blight (FHB), primarily caused by Fusarium pathogens, refer to are economically devastating diseases in wheat (*Triticum aestivum* L.) as well as other small grain cereals across the world (Bai and Shaner, 2004; Li X. et al., 2010; Ren et al., 2016). Fusarium seedling blight can cause extensive damage to growing seedlings or foot rot later during the growing season, leading to reduced emergence and crop establishment and consequently yield losses in wheat (Wiese, 1987; Antalová et al., 2020). Moreover, FSB can provide a pathogen source

for following FHB infection, creating reddish scabby spikes (Haigh et al., 2009; Li X. et al., 2010). Due to the global climate change and tillage management, FSB and FHB usually reach epidemic levels, causing huge yield losses across millions of hectares in global wheat production regions (Cheng et al., 2015; Liu et al., 2016). In addition, both FSB and FHB produce various mycotoxins during infection, with high toxicity, posing a threat to people as well as livestock (Pestka and Smolinski, 2005; Liu et al., 2012).

China is the largest producer and consumer of wheat (Shi and Ling, 2018). Cultivars play a major role in national wheat production, and developing and using resistant cultivars can confer protection to *Fusarium* pathogens. The analysis of the probable association between FSB and FHB can help develop new strategies to combat the *Fusarium* disease complex. Twelve Polish spring wheat cultivars and 18 spring wheat accessions from CIMMYT were examined for resistance to FSB and FHB by applying a highly aggressive fungal isolate, and no correlation was found between the two resistance types (Wisniewska and Busko, 2005). No significant correlation was also detected between FSB infection and FHB index and between FSB infection and DON content in a Wuhan/Nyubai doubled haploid (DH) wheat population (Somers et al., 2003; Tamburic-Ilincic et al., 2009). In subsequent research, QTLs for *Fusarium* resistance at seedling and spike stages were different, but further verification was required for various wheat populations (Tamburic-Ilincic et al., 2009). Comparatively, there have also been reports concerning the positive association between FSB and FHB resistance (Mesterhazy, 1987; Wu et al., 2005; Shin et al., 2014). Mesterhazy (1987) found a significant correlation between FSB and FHB resistance, and the most resistant genotypes at the seedling stage could yield the FHB resistant material with a large probability. By inoculation of wheat coleoptiles with *Fusarium graminearum* isolates, Wu et al. (2005) found a significant correlation between FSB and FHB resistance in the same genotype in the field. Using the clip-dipping inoculation method, Shin et al. (2014) found the remarkable correlation between the lesion length and Type II FHB resistance and suggested that the method for the evaluation of FSB resistance may provide a simple and feasible way for the early screening of FHB resistance in wheat.

Using linkage analysis, quite a few QTLs associated with FHB resistance were detected in 21 wheat chromosomes reported (Buerstmayr et al., 2009; Ma et al., 2020), with 7 FHB genes (*Fhb1-Fhb7*) being formally cataloged (Zhu et al., 2020). But FSB has not received much attention so far, and there have been very few studies on the QTLs for FSB resistance. A QTL on chromosome 5B controlling FSB resistance was identified in a DH population, and its linked marker WMC75 interpreted 13.8% of the phenotypic variation (Tamburic-Ilincic et al., 2009). Single major QTLs for FSB resistance, caused by *Microdochium nivale* and *Microdochium majus*, were detected on the chromosomes 1AL and 2BS, respectively (Ren et al., 2016).

Genome-wide association studies (GWAS) on the basis of linkage disequilibrium (LD) offer several advantages over linkage mapping, which has gained success in the analysis of different quantitative characteristics in wheat (Sapkota et al., 2019; Hu et al., 2020). For example, using 166 elite wheat varieties from

Yellow and Huai River Valleys Wheat District in China, 120 common loci were detected for their associations with grain yield, among which 78 were potentially new (Li et al., 2019). In our previous studies, five QTLs were identified for their consistent associations with FHB resistance in a natural population, among which the QTLs on 5AS, 5AL, and 7DS were possibly new (Zhu et al., 2020). However, GWAS to identify FSB in wheat has not been reported yet, and the molecular mechanisms for FSB remain poorly understood.

In the present study, we evaluated FSB resistance in Chinese elite wheat lines and then performed GWAS and QTL analyses. The study aimed to (1) identify wheat germplasms with FSB resistance that could be used as resistance donors in breeding and confirm the relationship between FSB and FHB resistance caused by *Fusarium* pathogens, (2) uncover novel FSB-resistant loci that could be used in molecular marker-assisted breeding. The findings provide an insight into the genetics of FSB response in Chinese cultivars, and the developed markers associated with the mapped QTLs may be used for breeding FSB resistance wheat.

MATERIALS AND METHODS

Plant Materials

A total of 240 common wheat cultivars or elite lines (**Supplementary Table 1**) were selected as the natural population to evaluate FSB resistance and perform GWAS analysis, as described in our previous study (Zhu et al., 2020). The population included 229 elite wheat cultivars (lines) developed in the main wheat-growing areas of China, covering 12 provinces with five agroecological systems, and could represent the current situation of breeding in China. The remaining 11 genotypes belonged to CIMMYT (10) and Australian (1). Seeds were harvested in the Wuhan Nanhu farm of Hubei Academy of Agricultural Sciences (N 30.28°, E 114.19°) during the cropping seasons in 2018–2019.

Phenotyping

Wheat coleoptiles at the seedling stage were inoculated with conidiospores using the previously described method (Li X. et al., 2010; Cheng et al., 2015) with minor alterations. For seedling inoculation, the concentration of the macroconidia suspension for the aggressive isolate *F. graminearum* Huanggang 1 (Zhu et al., 2020) was regulated to 5×10^5 spores/mL with sterilized distilled water. Forty full wheat seeds per cultivars (lines) were disinfected using 0.1% HgCl₂ for 1 min and then rinsed twice using sterilized distilled water. Sterilized seeds were placed on wet filter paper in Petri dishes and incubated at 20°C in the dark for 2 days. Then 20 seeds with steady growth were picked, transferred to a sterilized germination box (length, width, and height of 11.5, 11.5, and 9.8 cm, respectively) with three layers of wet filter papers, and kept in the dark at 20°C for 1 day. Top coleoptiles (2–3 mm) were dissected, and a 3-μL aliquot of the macroconidial suspension was injected into the slant side of the dissected seedlings. Inoculated seedlings were stored in the germination box in the dark at 20°C, in dark as previously mentioned. The brown lesions of diseased seedlings were measured at 7 day post

inoculation, and the lesion length was determined as previously described (Li X. et al., 2010). For each genotype, 20 wheat seedlings were examined each time, and the average value was used for subsequent analysis. The experiments were performed independently with triple replications.

Statistical Analysis

The *t*-tests, as well as Pearson's correlation analysis for independent samples, were conducted using IBM SPSS Statistics version 19.0 (IBM Corporation, Armonk, NY, United States). Histograms showing the distribution of the lesion length (cm) of 240 cultivars (lines) were made for each replicate with a script executed in R version 3.5.1.¹

Genotyping

Illumina 90 K SNP array genotyping was performed on 240 wheat accessions (Wang et al., 2014). Calling and filtering for SNPs, kinship, and population structure analysis have all been all elaborately explained in the previous research (Zhu et al., 2020). A total of 19,803 with MAF of >5 and <20% missing data of 22,922 polymorphic SNPs were employed for subsequent analysis (Zhu et al., 2020). Population structure analysis was performed via ADMIXTURE.² The population fell into three subgroups, basically based on geographic origin and pedigree (Zhu et al., 2020).

Genome-Wide Association Studies for Fusarium Seedling Blight Resistance

Associations between genotypic and phenotypic data were analyzed in Tassel v5.0. A kinship (K) + PCA model was used to perform the MLM analysis for controlling the background variation and eliminating spurious marker-trait associations (MTAs). R^2 exhibiting the variation explained by SNP was documented (Bradbury et al., 2007). SNPs with an adjusted- \log_{10} (*P*-value) of ≥ 3.0 were considered associated with FSB resistance. The remarkable loci in a minimum of two repetitions detected in the research were stable QTLs. Remarkable SNP markers in one linkage disequilibrium on the same chromosome represented one locus.

Kompetitive Allele-Specific PCR Assay

The SNP markers remarkably associated with FSB resistance were identified using GWAS and transformed into Kompetitive Allele-Specific PCR (KASP) markers to facilitate their application in MAS. The SNP contextual sequences were obtained at GrainGenes³ and the primers were designed by PolyMarker⁴ or the primer premier 5.0 (PREMIER Biosoft International, Palo Alto, CA, United States). Amplification was performed initially at 95°C for 15 min, 10 cycles of touchdown PCR (at 95°C for 20 s; an initial touchdown at 65°C, followed by a reduction of -1°C per cycle for 25 s), and the final 30 additional cycles for annealing (95°C for 10 s; 60°C for 60 s). Fluorescence signals were inspected

under the multifunctional microplate reader PHERAstarPlus (BMG LABTECH, Ortenberg, Germany) and determined via KlusterCaller (LGC Genomics, Teddington, United Kingdom).

Candidate Gene Analysis

To identify the candidate genes associated with typical SNPs, physical positions of markers before the chromosome name were introduced into Ensembl,⁵ and genes within a 2 Mb distance from typical SNPs were detected to assess their candidacy for FSB resistance. The transcript IDs of all these genes were obtained from wheat sequences (Alaux et al., 2018). We used another publicly available database, expVIP,⁶ to obtain the expression profiles of all these genes in wheat seedling coleoptile organs infected by *Fusarium* spp. (Ma et al., 2014; Powell et al., 2017). To visualize the expression profiles, heat maps were drawn using TBtools (Chen et al., 2018) from differently expressed genes, with the absolute value of \log_2 fold change of ≥ 1 or ≤ -1 at either time point. Up-regulated genes in resistant varieties associated with disease resistance annotated by RefSeq Annotation v1.1 (Appels et al., 2018) were identified as candidate genes. The candidate genes were further used for searing the sequences with high similarity via NCBI, combined with a basic local alignment search tool (BLAST).⁷

RESULTS

The Evaluation of Fusarium Seedling Blight Resistance

The assessment of FSB resistance in 240 wheat accessions showed a lesion length within the range of 0.075–2.896, with a normal distribution in three replications (Table 1 and Figure 1). Pearson's correlation coefficients among three replications ranged from 0.535 to 0.577 with a significant difference ($P < 0.01$), and the average values were significantly associated with repeats, with correlation coefficients of 0.826, 0.857, and 0.828 for each repeat (Table 1). Further analysis indicated that 23 accessions, including the elite cultivars Zhoumai 17, Yanzhan 4110, Yunong 035, Jimai 38, Yumai 69, Jingfumai 1, Zhoumai 16, Yannong 24, and Yangmai 11 showed an average lesion length of less than 0.6, with a potential for breeding for FSB resistance in wheat. There were 105, 71, 29, and 12 accessions showing the average lesion lengths within the ranges of 0.6–1.0, 1.01–1.40, 1.41–1.80, and >1.80, respectively. Representative accessions with different grades of resistance to FSB are represented in Table 2.

The Correlation Between Fusarium Seedling Blight and Fusarium Head Blight

Correlation coefficients were determined based on FSB infection values in the research and BLUE values of FHB indices, calculated

¹ www.r-project.org

² <http://software.genetics.ucla.edu/admixture>

³ <https://wheat.pw.usda.gov/GG3/>

⁴ <http://polymarker.tgac.ac.uk>

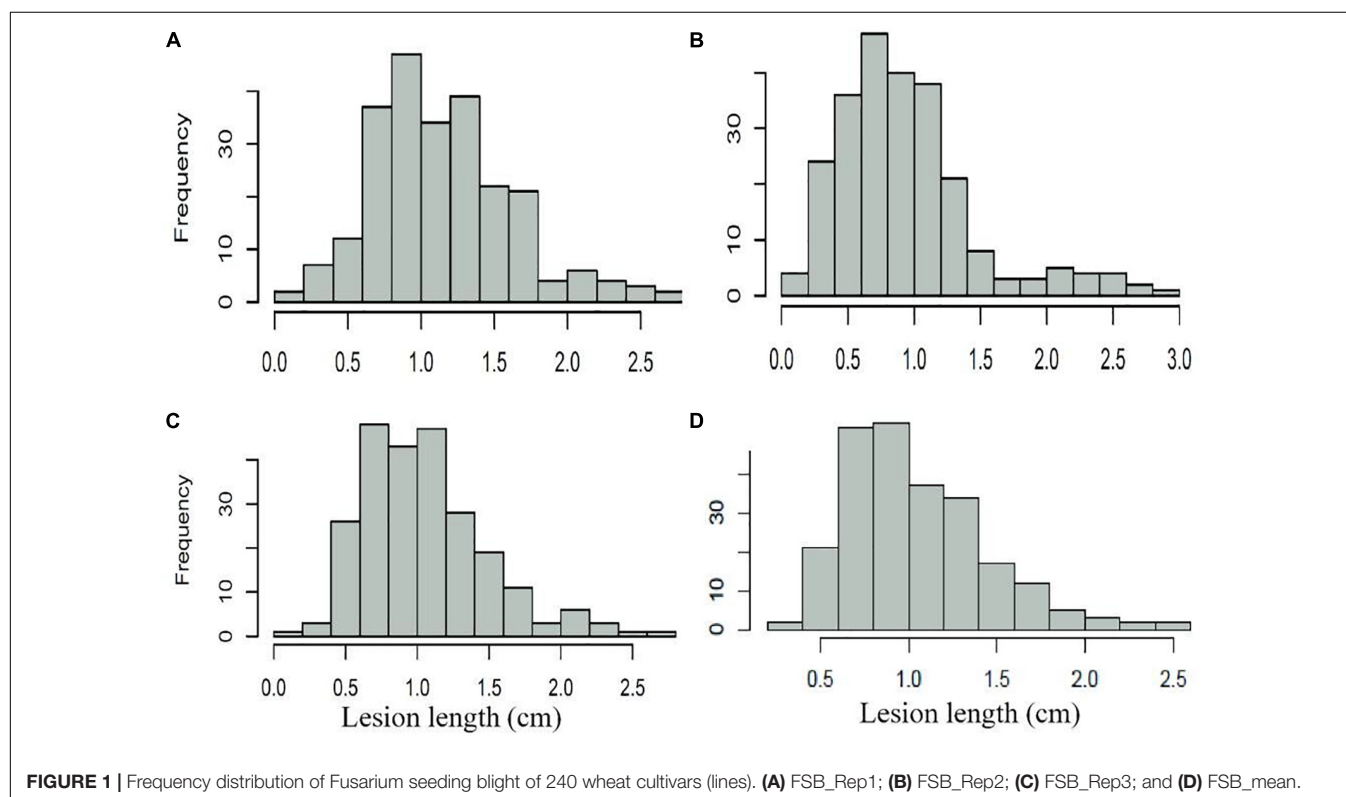
⁵ https://urgi.versailles.inra.fr/gb2/gbrowse/wheat_survey_sequence_annotation

⁶ <http://www.wheat-expression.com/>

⁷ <http://blast.ncbi.nlm.nih.gov/Blast.cgi>

TABLE 1 | Descriptive statistics and correlation coefficients of Fusarium seedling blight and Fusarium head blight of 240 wheat cultivars (lines).

| | T-TEST | | | | | Correlations | | | |
|--------------------------|--------|-------|-------|--------|-----------|--------------|----------|----------|----------|
| | Mean | Min | Max | SD | Std Error | FSB_Rep1 | FSB_Rep2 | FSB_Rep3 | FSB_Mean |
| FSB_Rep1 | 1.159 | 0.139 | 2.769 | 0.478 | 0.031 | | | | |
| FSB_Rep2 | 0.946 | 0.075 | 2.896 | 0.531 | 0.034 | 0.535** | | | |
| FSB_Rep3 | 1.052 | 0.165 | 2.735 | 0.437 | 0.028 | 0.545** | 0.577** | | |
| FSB_Mean | 1.049 | 0.346 | 2.510 | 0.404 | 0.026 | 0.826** | 0.857** | 0.828** | |
| BLUE of FHB ^a | 46.66 | 5.00 | 89.00 | 16.571 | 1.070 | -0.239** | -0.160* | -0.273** | -0.263** |

Significant at $P < 0.01$.*Significant at $P < 0.05$.^aData from our previous study (Zhu et al., 2020).FIGURE 1** | Frequency distribution of Fusarium seedling blight of 240 wheat cultivars (lines). (A) FSB_Rep1; (B) FSB_Rep2; (C) FSB_Rep3; and (D) FSB_mean.

within 4 years, according to the results of our previous study (Zhu et al., 2020) using the same population. The average FSB lesion length was negatively correlated with the BLUE FHB index across the population, although a low coefficient of $R = -0.263$ was determined (Table 1). The most notable cultivar Sumai3 and its derivative Ning7840 with a high FHB resistance showed quite low resistance to FSB in this assay. Conversely, the FHB susceptible cultivars Zhengyumai 9987 and Zhoumai 17 (Zhu et al., 2020) showed relatively high resistance to FSB. However, some accessions such as Jingfumai 1 and Yangmai 11 had relatively high resistance to both FSB and FHB simultaneously (Zhu et al., 2020).

Marker-Trait Association Analysis

Six QTLs on chromosome arms 1DL, 3AS, 3BL, 6BL, 7AL, and Un, designated as *Qfsb.hbaas-1DL*, *Qfsb.hbaas-3AS*,

Qfsb.hbaas-3BL, *Qfsb.hbaas-6BL*, *Qfsb.hbaas-7AL*, and *Qfsb.hbaas-un*, respectively, were significant in a minimum of two repetitions, interpreting phenotypic variation of 4.83–7.53% (Table 3 and Figure 2). Representative significant markers for these QTLs were IWB41243, IWB64668, IWB3107, IWA3221, IWB41907, and IWB36312, respectively. For the goal of identifying minor QTL, this study was underpowered because of small population size which results in not seeing high signals.

Of the 240 genotypes, 12, 180, 130, 204, 220, and 217 possessed the resistance alleles *Qfsb.hbaas-1DL*, *Qfsb.hbaas-3AS*, *Qfsb.hbaas-3BL*, *Qfsb.hbaas-6BL*, *Qfsb.hbaas-7AL*, and *Qfsb.hbaas-un*, respectively, on the basis of marker analysis (Table 4 and Supplementary Table 1). The mean FSB lesion length in accessions with favorable *Qfsb.hbaas-1DL* alleles was 21.4% shorter than with unfavorable alleles. Discrepancies

TABLE 2 | Materials with different resistance levels to Fusarium seedling blight (FSB) (only representative materials are shown).

| Mean of lesion length (cm) | Number | Representative cultivars (lines) |
|----------------------------|--------|--|
| ≤0.60 | 23 | Zhengyumai 9987, Zhoumai 17, Yanzhan 4110, Yunong 035, Lianmai 2, Xikemai 4, Yumai 70–36, Zhongmai 1, Yan 2415, Luohan 2, Luomai 21, Jimai 38, Yumai 69, Jingfumai 1, Zhoumai 16, Yannong 24, and Yangmai 11 |
| 0.60–1.00 | 105 | Xinmai 20, SYN1, Kenong 199, Zhongyu 10, Yangmai 22, Ningmai 13, Chuanmai 42, Zhengmai 366, Xinong 9871, Yangmai 13, Liangxing 99, Xinong 979, Jimai 22, Ocoroni, Mianmai 37, Xiaoyan 22, Yangmai 17, Ningmai 9, Hengguan 35, and Emai 27 |
| 1.01–1.40 | 71 | Lantian 18, Jingdong 17, Emai 23, Yangmai 12, Xinmai 11, Mayoor, and Huaimai 20 Zhongmai 9, Zhengmai 9023, Jimai 20, Ningdong 10, 04 Zhong 36, Aikang 58, Chuanmai 50, Wuhan 1, Zhenmai 168, Ningchun 43, Jingmai 103, Lumai 21, Ningmai 16, Pingan 6, and Emai 580 |
| 1.41–1.80 | 29 | Yumai 48, Emai 12, Een 6, Chuanmai 51, Yangmai 16, Lantian 23, Ningdong 11, Yangmai 158, Ningmai 8, Emai 18, Lunxuan 987, Xiangmai 25, Jingdong 8, and Xiaoyan 6 |
| > 1.80 | 12 | Xinong 88, Jingzhou 66, Ningmai 11, Ning 7840, Zhongnong 2, Xiangmai 55, Jining 16, Een 5, Een 1, Sumai 3, and Gamanya |

TABLE 3 | Loci significantly associated with FSB resistance in at least two environments in the 240 wheat cultivars (lines) using the mixed linear model (MLM) model in Tassel v5.0.

| QTL | Marker ^a | Variant ^b | Chr ^c | Position (Mb) ^d | Environment | P-value | R ² (%) ^e |
|-----------------------|---------------------|----------------------|------------------|----------------------------|----------------|-----------------------------|---------------------------------|
| <i>Qfsb.hbaas-1DL</i> | <i>IWB41243</i> | <u>A</u> /G | 1DL | 458.9 | Rep2/Mean | 6.36E-04/7.37E-04 | 5.74/5.33 |
| <i>Qfsb.hbaas-3AS</i> | <i>IWB64668</i> | <u>T</u> /G | 3AS | 176.6 | Rep1/Mean | 4.57 E-04/8.16 E-04 | 5.12/4.83 |
| <i>Qfsb.hbaas-3BL</i> | <i>IWB3107</i> | <u>G</u> /A | 3BL | 723.0 | Rep1/Mean | 3.24 E-04/2.14 E-04 | 5.47/6.29 |
| <i>Qfsb.hbaas-6BL</i> | <i>IWA3221</i> | <u>C</u> /T | 6BL | 668.0 | Rep1/Rep3/Mean | 6.34E-04/6.39E-04/1.21 E-04 | 5.07/5.20/6.50 |
| <i>Qfsb.hbaas-7AL</i> | <i>IWB41907</i> | <u>G</u> /A | 7AL | 724.1 | Rep1/Mean | 5.86 E-05/6.99 E-05 | 7.00/7.53 |
| <i>Qfsb.hbaas-un</i> | <i>IWB36312</i> | <u>A</u> /C | Un | 32.2 | Rep2, Mean | 2.51 E-04/9.40 E-05 | 6.18/6.74 |

^aRepresentative markers showing the strongest association with the FSB resistance locus.

^bFavorable allele is underlined.

^cChr, chromosome.

^dPhysical positions based on the Chinese Spring reference genome sequences from the International Wheat Genome Sequencing Consortium (IWGSC, <http://www.wheatgenome.org>).

^ePercentage of phenotypic variance explained.

between *Qfsb.hbaas-6BL* and *Qfsb.hbaas-7AL* were much greater (31.7 and 36.3%, respectively). In *Qfsb.hbaas-3AS*, *Qfsb.hbaas-3BL*, and *Qfsb.hbaas-un*, the FSB lesion lengths were reduced by 17.4, 13.8, and 8.8%, respectively (Table 4).

The Relationship Between the Fusarium Seedling Blight Lesion Length and the Number of Favorable Alleles

To examine the pyramiding effects of favorable alleles of various QTLs, we analyzed the number of favorable alleles in 6 mapped loci per accession. Favorable alleles were 0–5. Linear regression ($r^2 = 0.872$) revealed the correlation between disease severity and the number of favorable alleles (Figure 3 and Supplementary Table 2). Accessions including a larger number of favorable alleles, such as Zhoumai16 (5), Xikemai4 (4), and Zhoumai17 (4), exhibited strong FSB resistance. Conversely, Yang 07–15, with no favorable alleles, exhibited low FSB resistance (Supplementary Table 1).

Development of Kompetitive Allele-Specific PCR Markers for Quantitative Trait Loci Underlying Resistance to Fusarium Seedling Blight

The SNPs (IWB41243, IWB64668, IWB3107, and IWA3221), associated with *Qfsb.hbaas-1DL*, *Qfsb.hbaas-3AS*,

Qfsb.hbaas-3BL, and *Qfsb.hbaas-6BL*, respectively, were successfully used to develop KASP markers (Table 5). All 240 wheat accessions were genotyped by these KASP markers. The results demonstrated that the genotypes from the KASP test were identical to the chip assay with low-frequency oscillations (2.5, 5.0, 3.3, and 3.8% for each marker, respectively).

The Prediction of Candidate Genes

A total of 291 candidate genes were located within the candidate regions. Combined with transcriptome data from public databases (Ma et al., 2014; Powell et al., 2017), 57 genes were differently expressed in wheat seedling coleoptile organs after infection by *Fusarium* spp. (Figure 4 and Supplementary Table 3). Among them, eight unique annotated genes involved in plant disease resistance were considered as candidates for mapping QTLs (Table 6). Two genes encoding the disease resistance protein RPM1 and receptor-like protein kinase were identified as candidates for *Qfsb.hbaas-1DL*. A gene encoding L-type lectin receptor kinase might contribute to FSB resistance for *Qfsb.hbaas-3AS*. A gene encoding MADS-box protein was considered as a candidate for *Qfsb.hbaas-6BL*. For *Qfsb.hbaas-7AL*, a gene encoding NAC domain-containing protein was identified. Three genes encoding serine/threonine kinase-like protein, HCBT-like defense response protein, and subtilisin-like protease might contribute to FSB resistance for *Qfsb.hbaas-un*.

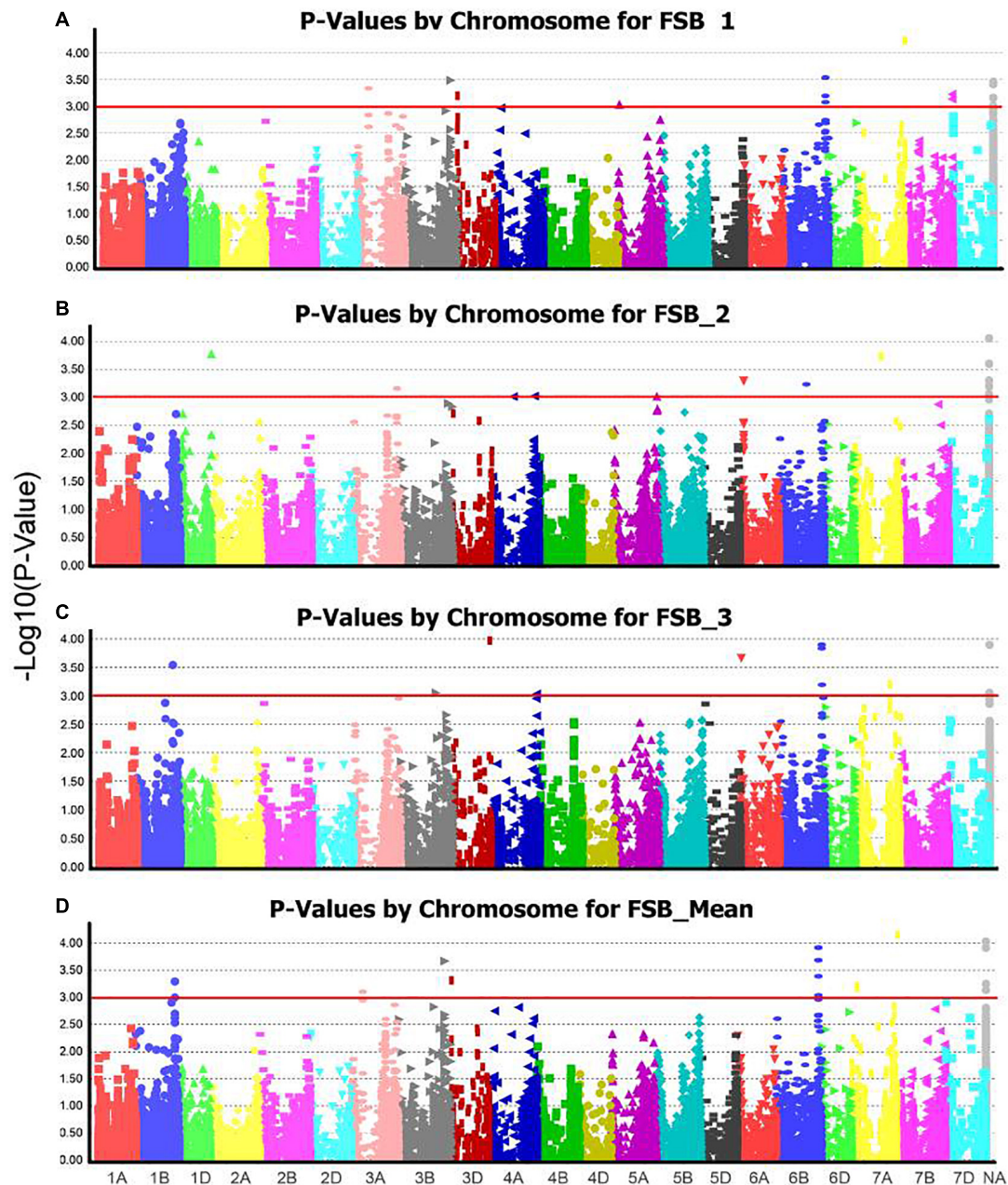


FIGURE 2 | Manhattan plots from genome-wide association scan for Fusarium seedling blight (FSB) severities among 240 wheat accessions in (A) FSB_Rep1, (B) FSB_Rep2, (C) FSB_Rep3, and (D) FSB_mean. Dashed red horizontal line is the significant threshold level.

DISCUSSION

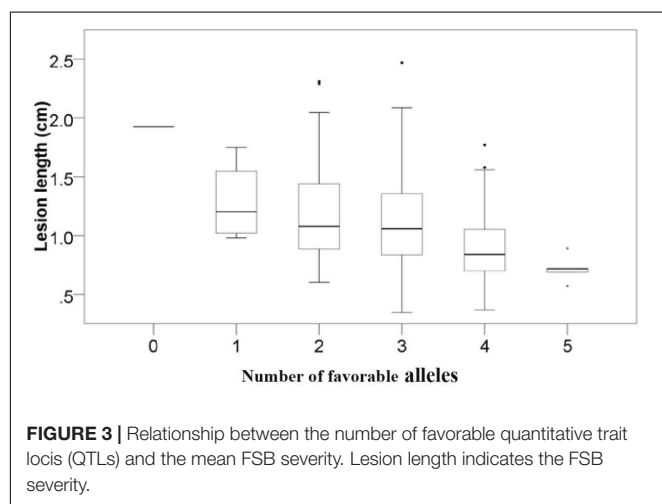
Breeding the cultivars with resistance to FSB and FHB offers an efficient way to control complex diseases and decrease yield losses or mycotoxin occurrence in agricultural products. Combining the two resistance traits in one elite wheat cultivar is challenging due to its exposure to regulated independent genetic loci and also the restricted size of resistant germplasm in the natural environment (Tamburic-Illincic et al., 2009). In this study, 229 elite Chinese wheat cultivars and lines, which represent the

genetic diversity in newly assembled accessions in China (Jia et al., 2020), were investigated. A total of 54 cultivars reached the maximum annual acreage of 1×10^5 ha during 2000–2016, and quite a few cultivars such as Liangxing 99, Zhoumai 18, Jimai 22, Zhoumai 22, Zhengmai 9023, and Aikang 58, have been used as founder parents in breeding programs (Jia et al., 2020). Among them, 23 accessions showed a relatively high-level resistance with an average lesion length of less than 0.6 cm, exhibiting the potential for breeding for FSB resistance in wheat. Despite the negative correlation between FSB infection and FHB index in

TABLE 4 | T-tests for differences in Fusarium seedling blight between two groups of wheat accessions with contrasting resistance or susceptibility alleles for quantitative trait loci (QTL) on chromosomes 1D, 3A, 3B, 6B, 7A, and Un.

| QTL | Present/Absent [#] | Number | Rep1 | Rep2 | Rep3 | Mean |
|----------------------|-----------------------------|--------|-------|-------|-------|-------|
| <i>Qfsb.hbaas-1D</i> | Present | 12 | 1.02a | 0.57a | 0.78a | 0.79a |
| | Absent | 223 | 1.15a | 0.95b | 1.06b | 1.06b |
| <i>Qfsb.hbaas-3A</i> | Present | 180 | 1.07A | 0.89A | 1.03a | 1.00A |
| | Absent | 58 | 1.39B | 1.11B | 1.14a | 1.21B |
| <i>Qfsb.hbaas-3B</i> | Present | 130 | 1.05A | 0.84A | 0.94A | 0.94A |
| | Absent | 108 | 1.26B | 1.06B | 1.17B | 1.16B |
| <i>Qfsb.hbaas-6B</i> | Present | 204 | 1.08A | 0.88A | 1.00A | 0.99A |
| | Absent | 31 | 1.56B | 1.40B | 1.41B | 1.45B |
| <i>Qfsb.hbaas-7A</i> | Present | 220 | 1.10A | 0.89A | 1.02A | 1.00A |
| | Absent | 15 | 1.69B | 1.48B | 1.54B | 1.57B |
| <i>Qfsb.hbaas-un</i> | Present | 217 | 1.15a | 0.92a | 1.05a | 1.04a |
| | Absent | 19 | 1.15a | 1.14a | 1.14a | 1.14a |

[#]The QTL present superior effect (present) or inferior effect (absent), A and B represent significant at $P < 0.01$, a and b represent significant at $P < 0.05$.



the population (Table 1), combined with our previous research results (Zhu et al., 2020), we also found that some cultivars such as Jingfumai 1 and Yangmai 11 had relatively great resistance to both FSB and FHB simultaneously. Jingfumai 1 and Yangmai 11 both bred in the Middle and Lower Yangtze River Valleys were red-grained spring wheat and high resistant to pre-harvest sprouting. The spike length of the two lines was 8.0 and 8.4 cm, and the spikelet number was 18.3 and 17.2, respectively. The research conformed to the findings reported by Ren et al. (2015), who found that the FSB-resistant cultivar Petrus was simultaneously resistant to FHB. These lines were good parent candidates for future crosses in breeding for *Fusarium* seedling resistance and head blight resistance in wheat.

The correlation analysis between FSB and FHB resistance has been reported in previous studies (Mesterhazy, 1987; Ruckebauer et al., 2001; Gosman et al., 2005; Tamburic-Ilincic et al., 2009; Shin et al., 2014). Few studies revealed a positive association between FSB and FHB resistance. Shin et al. (2014) reported the significant correlation coefficients between the

TABLE 5 | Primer sequences of single nucleotide polymorphism (SNP) markers for validation in wheat lines by Kompetitive Allele-Specific PCR (KASP) assay.

| QTL | Primer | Sequence (5'-3') |
|----------------------|---------|---------------------------|
| <i>Qfsb.hbaas-1D</i> | P41243A | CCACCTTTCAACTCGCTCA |
| | P41243B | CCACCTTTCAACTCGCTCG |
| | P41243C | CTCACTTCTTCTAGAACAAATCGAA |
| <i>Qfsb.hbaas-3A</i> | P64668A | TGCAATCTTGGACAACATCAT |
| | P64668B | TGCAATCTTGGACAACATCAG |
| | P64668C | GTGCTTTGTCAACAACAGATGC |
| <i>Qfsb.hbaas-3B</i> | P3107A | GGTCGCATCAGGAAGAGCA |
| | P3107B | GGTCGCATCAGGAAGAGCG |
| | P3107C | TTCTTCCTTTACAGACTCTTCAGC |
| <i>Qfsb.hbaas-6B</i> | P3221A | GTTTTGTGGCTGCGGGT |
| | P3221B | GTTTTGTGGCTGCGGGC |
| | P3221C | TTCTTCCTTTACAGACTCTTCAGC |

A Primer labeled with FAM: GAAGGTGACCAAGTTTCATGCT.

B Primer labeled with HEX: GAAGGTGCGAGTCAACGGATT.

lesion lengths and Type II resistance to FSB and FHB, but the number of samples was not very large and only 29 Korean winter wheat cultivars were chosen in trials. The CIMMYT spring wheat line LSP2 was proved to have a high susceptibility to FSB and resistance to FHB, caused by *Fusarium* spp. (Ren et al., 2016). The widely planted British winter wheat cultivar Rialto was highly resistant to FSB, caused by *Microdochium* spp., while some reports revealed its high susceptibility to FHB (Srinivasachary et al., 2008). Some wheat cultivars, including Chinese local cultivars Wangshuibai and Sumai3, were highly resistant to FHB, and high susceptibility to FSB was also found (Mesterhazy, 1987; Wu et al., 2005; Li X. et al., 2010). Using Sumai3 and Falat as the cultivars resistant and susceptible to FHB, respectively, Sorahinobar et al. (2016) observed little correspondence between wheat seedling tolerance to *F. graminearum* crude extract and resistance to FHB. In both our previous (Zhu et al., 2020) and present studies, Sumai3 showed an FSB-susceptible reaction while exhibiting FHB resistance in response to *Fusarium* spp. Negative correlations, albeit low, between FSB and FHB resistance were observed in the present study. These results also agreed with the findings published by Bruins et al. (1993); Ruckebauer et al. (2001); Gosman et al. (2005), and Tamburic-Ilincic et al. (2009), who discovered that greenhouse experiments in seedling cannot be used when selecting for FHB resistance.

Two transgenic wheat lines expressing two anti-fungal peptides exhibited enhanced resistance to FSB and FHB, while FHB resistance could be detected in the other five lines (Liu et al., 2012). Transgenic wheat overexpressing an *A. thaliana* *NPR1* gene could increase the severity of FSB, although FHB resistance increased simultaneously (Gao et al., 2013). Li X. et al. (2010) firstly reported a close association between FHB and FSB resistance in wheat using distinct molecular profiles for disease-associated gene expression and suggested that there may be two resistance mechanisms in wheat spikes and seedlings in response to FHB pathogens. Some studies have also shown different QTLs for resistance to FSB and FHB (Tamburic-Ilincic et al., 2009; Ren et al., 2016). In our previous study, five QTL on chromosome

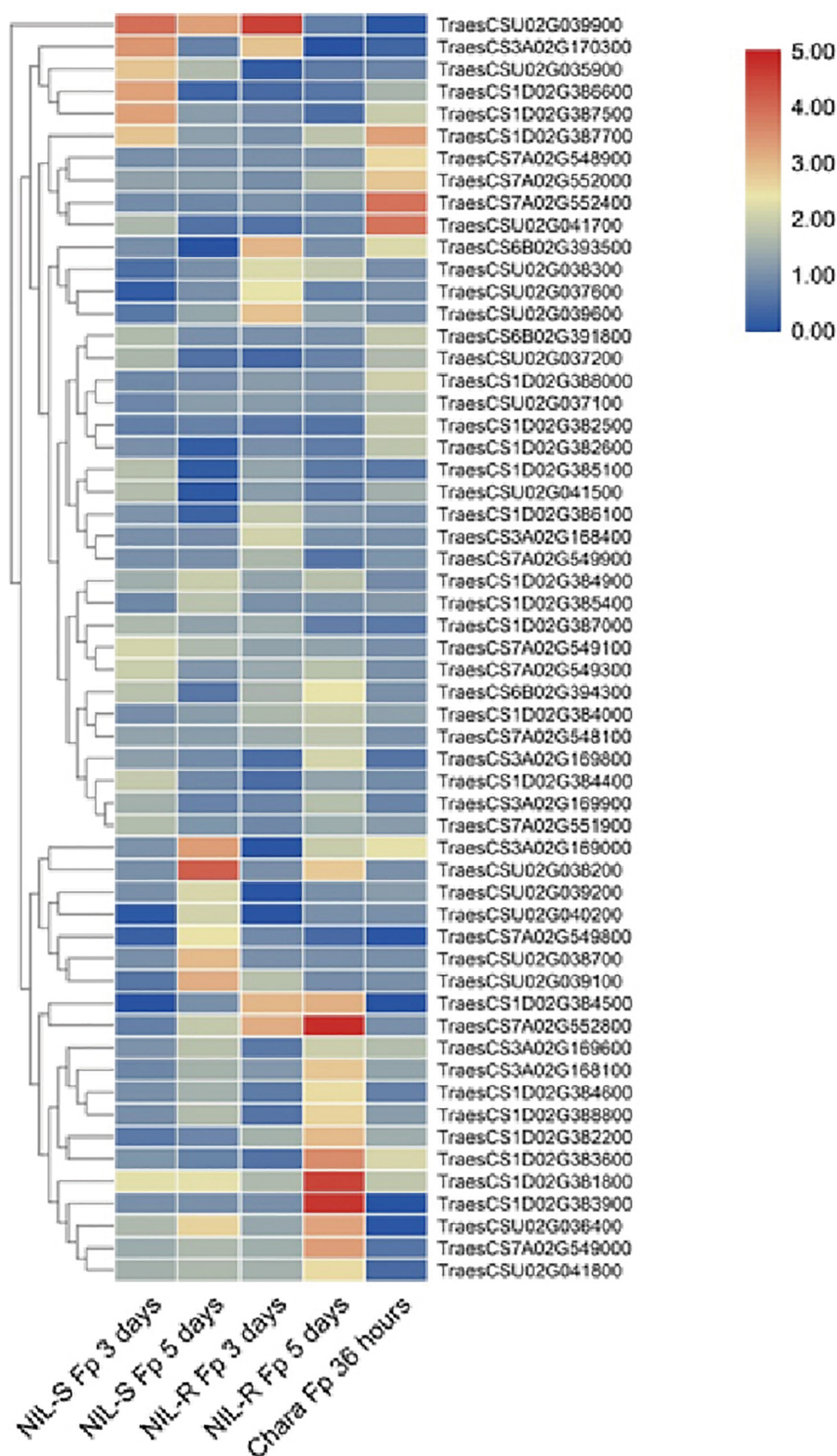


FIGURE 4 | Clustering heatmap for differentially expressed genes infected by *Fusarium* spp. compared to those in the mock. S-NIL1-Fp 3 days, susceptible isolines infected by *Fusarium* spp. after 3 days; S-NIL1-Fp 5 days, susceptible isolines infected by *Fusarium* spp. after 5 days; R-NIL1-Fp 3 days, resistant isolines infected by *Fusarium* spp. after 3 days; R-NIL1-Fp 5 days, resistant isolines infected by *Fusarium* spp. after 5 days; Chara-Fp 36 h, wheat culture “Chara” infected by *Fusarium* spp. after 36 h.

arms 1AS, 2DL, 5AS, 5AL, and 7DS were associated with FHB resistance, explaining 5.4–10.3% of phenotypic variation (Zhu et al., 2020). Using the same population, we identified six entirely different QTL on chromosome arms 1DL, 3AS, 3BL, 6BL, 7AL, and Un, interpreting phenotypic variation of 4.83–7.53% (Table 3 and Figure 2). Different regions suggested the differences in QTLs for resistance to FSB and FHB and resistance to FSB and FHB is probably independent. Thus, due to gene recombination, a few accessions in this research exhibited Fusarium resistance in seedling and head.

There might be two main reasons why discrepant alterations in genes or QTLs could be associated with resistance to both FSB and FHB. The first reason is the infection time; FSB infection occurs during the seedling growth, whereas FHB infection occurs during the flowering stage. It is well known that different genes can be involved in the resistance of host plants to one disease in various stages of plant development (Li H. B. et al., 2010). The second reason is the infection of different organs by the two diseases. Miedaner (1997) put forward the idea that complex interactions can occur between the resistance to diseases across different plant growth stages, plant organs, or host genotypes. We suggested that the mechanisms and genes involved in resistance to Fusarium during seedling growth and spike formation are possibly different, and separate screening is essential to evaluate the resistance to FSB and FHB, caused by *Fusarium* in breeding programs.

The discovery of novel genes or QTLs is a constant challenge and extremely important in wheat breeding. Many QTLs associated with FHB resistance have been detected (Zhu et al., 2020), whereas there have been very few studies on QTLs for resistance to FSB. Using a Wuhan/Nyubai doubled haploid (DH) wheat population, merely one QTL, controlling FSB resistance, detected on chromosome 5B and marker WMC75, could interpret 13.8% phenotypic variation in the trait (Tamburic-Illincic et al., 2009). Using the Rialto/LSP2 DH population, a single major QTL conferring FSB resistance to *Microdochium majus* was found on the chromosome 1AL in all four experiments, accounting for 32.5–56.6% of the phenotypic variation; a significant QTL conferring FSB resistance to *Microdochium nivale* was discovered on chromosomes 2BS and explained 29.3–55.0% of the phenotypic variation (Ren et al., 2016). In this

research, we detected six QTLs on chromosome arms 1DL, 3AS, 3BL, 6BL, 7AL, and Un that were significant for resistance to FSB and previously uncharacterized in wheat, and therefore they were likely to be novel QTLs for FSB resistance. The average lesion length dramatically decreased when the number of favorable alleles increased (Figure 3). This finding suggested the prospective role of these QTLs in FSB resistance. But some QTLs such as *Qfsb.hbaas-1D* and *Qfsb.hbaas-7A* have very tiny minor allele frequencies, they may not be real signals. To validate the real effects of these 6 mapped QTLs, typical significant markers should be examined using various bi-parental and natural populations.

Combined with the analysis of transcriptome data, we identified eight unique annotated genes involved in plant disease resistance in wheat in IWGSC RefSeq v1.1, which were linked to the six QTLs (Table 6). The RPM1 is a CC-NB-LRR protein that conferred resistance against *Pseudomonas syringae* pv. *maculicola* 1 (Mackey et al., 2002; Su et al., 2017), and TaRPM1 might play a key part in the wheat innate immune response to the infection caused by the powdery mildew pathogen (Nie and Ji, 2019). Receptor-like protein kinases, which are the largest gene family in plants, play essential roles in combating infection caused by pathogens (Liu et al., 2017). TaCRK2, a novel receptor-like kinase gene, plays a positive role in resistance to leaf rust in wheat through the regulation of the HR cell death process induced by *P. trititica* (Gu et al., 2020). L-type lectin receptor kinases are omnipresent in plants and play an important role in the initiation of innate immunity (Wang and Bouwmeester, 2017). An L-type lectin receptor kinase in *Haynaldia villosa* conferred powdery mildew resistance in wheat (Wang et al., 2018). Moreover, MIKC-type MADS-box genes exhibited new expression patterns in response to biotic stress (Schilling et al., 2020). The NAC protein constituting the most important plant transcription factors could enhance resistance to Fusarium head blight, as well as stripe rust (Ning et al., 2010; Perochon et al., 2019). Serine/threonine kinase, one of the largest protein kinase gene families, could confer resistance to powdery mildew and stripe rust in wheat (Cao et al., 2011; Gou et al., 2015). HCBT-like defense response protein, which was rapidly and transiently expressed after being induced by the pathogen, plays an essential role in fungal pathogen resistance (Brooks et al., 2002). The

TABLE 6 | Candidate genes for Fusarium seedling blight resistance.

| Gene ID | Chr ^a | Position ^b (Mb) | Predicted function ^c | Identity (%) | Orthologous gene |
|--------------------|------------------|----------------------------|--------------------------------------|--------------|------------------|
| TraesCS1D02G388800 | 1D | 460.85 | Disease resistance protein RPM1 | 99.40 | LOC109754777 |
| TraesCS1D02G381800 | 1D | 457.07 | Receptor-like protein kinase | 100 | LOC109748921 |
| TraesCS3A02G169600 | 3A | 177.70 | L-type lectin receptor kinase | 98.82 | LOC109777203 |
| TraesCS6B02G391800 | 6B | 666.77 | MADS-box protein | 99.36 | LOC119321270 |
| TraesCS7A02G549000 | 7A | 723.26 | NAC domain-containing protein | 90.58 | LOC109760823 |
| TraesCSU02G041800 | Un | 34.16 | Serine/threonine kinase-like protein | 83.01 | LOC109786647 |
| TraesCSU02G041700 | Un | 34.14 | HCBT-like defense response protein | 92.54 | LOC109754238 |
| TraesCSU02G039600 | Un | 32.13 | Subtilisin-like protease | 87.27 | LOC109786891 |

^aChr, chromosome.

^bGene annotations were referred to IWGSC Ref Seq annotation v1.1 (IWGSC, <http://www.wheatgenome.org/>).

^cThe sequences of *T. aestivum* gene were blasted in the NCBI (<http://www.ncbi.nlm.nih.gov/>), databases to identify putative gene functions.

subtilisin-like protease is associated with pathogenicity in fungi and plays an important role in resistance to leaf rust in wheat (Fan et al., 2016).

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

YL, DM, and CG guided the design of the experiment. YL, GZ, ZZ, LC, HN, WH, HT, JS, and YZ directed the data analysis. YL and GZ conducted data analysis and wrote the manuscript. DM and CG supervised the experiment and confirmed the manuscript. YL was the guarantor of this work, so she could

have full access to all the data in the research and responsible for the integrity of the data and the accuracy of the data analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by Hubei Provincial Special Project of Central Government Guides Local Science and Technology Development (2020ZYYD011) and the Hubei Key Laboratory of Food Crop Germplasm and Genetic Improvement (2018zjj09).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.777494/full#supplementary-material>

REFERENCES

- Alaux, M., Rogers, J., Letellier, T., Flores, R., Alfama, F., Pommier, C., et al. (2018). Linking the International Wheat Genome Sequencing Consortium bread wheat reference genome sequence to wheat genetic and phenomic data. *Genome Biol.* 19:111. doi: 10.1186/s13059-018-1491-4
- Antalová, Z., Bleša, D., Martinek, P., and Matušinsky, P. (2020). Transcriptional analysis of wheat seedlings inoculated with *Fusarium culmorum* under continual exposure to disease defence inducers. *PLoS One* 15:e0224413. doi: 10.1371/journal.pone.0224413
- 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. doi: 10.1126/science.aar7191
- Bai, G., and Shaner, G. (2004). Management and resistance in wheat and barley to *Fusarium* head blight. *Annu. Rev. Phytopathol.* 42, 135–161. doi: 10.1146/annurev.phyto.42.040803.140340
- 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
- Brooks, S. A., Huang, L., Gill, B. S., and Fellers, J. P. (2002). Analysis of 106 kb of contiguous DNA sequence from the D genome of wheat reveals high gene density and a complex arrangement of genes related to disease resistance. *Genome* 45, 963–972. doi: 10.1139/g02-049
- Bruins, M. B., Karsai, M., Schepers, I. J., and Snijders, C. H. A. (1993). Phytotoxicity of deoxynivalenol to wheat tissue with regard to in vitro selection for *Fusarium* head blight resistance. *Plant Sci.* 94, 195–206. doi: 10.1016/0168-9452(93)90020-Z
- Buerstmayr, H., Ban, T., Anderson, J. A., Buerstmayr, H., Ban, T., and Anderson, J. A. (2009). QTL mapping and marker-assisted selection for *Fusarium* head blight resistance in wheat: a review. *Plant Breed.* 128, 1–26. doi: 10.1556/crc.36.2008.suppl.b.1
- Cao, A., Xing, L., Wang, X., Yang, X., Wang, W., Sun, Y., 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
- Chen, C., Rui, X., Hao, C., and He, Y. (2018). TBtools, a toolkit for biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv [preprint]* 289660. doi: 10.1101/289660
- Cheng, W., Song, X. S., Li, H. P., Cao, L. H., Sun, K., and Qiu, X. L. (2015). Host-induced gene silencing of an essential chitin synthase gene confers durable resistance to *Fusarium* head blight and seedling blight in wheat. *Plant Biotechnol. J.* 13, 1335–1345. doi: 10.1111/pbi.12352
- Fan, T., Bykova, N. V., Rampitsch, C., and Xing, T. (2016). Identification and characterization of a serine protease from wheat leaves. *Eur. J. Plant Pathol.* 146, 293–304. doi: 10.1007/s10658-016-0914-x
- Gao, C. S., Kou, X. J., Li, H. P., Zhang, J. B., Saad, A. S. I., and Liao, Y. C. (2013). Inverse effects of Arabidopsis NPR1 gene on fusarium seedling blight and fusarium head blight in transgenic wheat. *Plant Pathol.* 62, 383–392. doi: 10.1111/j.1365-3059.2012.02656.x
- Gosman, N., Chandler, E., Thomsett, M., Draeger, R., and Nicholson, P. (2005). Analysis of the relationship between parameters of resistance to *Fusarium* head blight and in vitro tolerance to deoxynivalenol of the winter wheat cultivar WEK0609H. *Eur. J. Plant Pathol.* 111, 57–66. doi: 10.1007/s10658-004-2733-8
- Gou, J. Y., Li, K., Wu, K., Wang, X. D., Lin, H. Q., Cantu, D., et al. (2015). Wheat stripe rust resistance protein WKS1 reduces the ability of the thylakoid-associated ascorbate peroxidase to detoxify reactive oxygen species. *Plant Cell.* 27, 1755–1770. doi: 10.1105/TPC.114.134296
- Gu, J., Sun, J., Liu, N., Sun, X., Liu, C., Wu, L., et al. (2020). A novel cysteine-rich receptor-like kinase gene, *TaCRK2*, contributes to leaf rust resistance in wheat. *Mol. Plant Pathol.* 21, 732–746. doi: 10.1111/mpp.12929
- Haigh, I. M., Jenkinson, P., and Hare, M. C. (2009). The effect of a mixture of seed-borne *Microdochium nivale* var. *majus* and *Microdochium nivale* var. *nivale* infection on *Fusarium* seedling blight severity and subsequent stem colonisation and growth of winter wheat in pot experiments. *Eur. J. Plant Pathol.* 124, 65–73. doi: 10.1007/s10658-008-9393-z
- Hu, W., Gao, D., Wu, H., and Liu, J. (2020). Genome-wide association mapping revealed syntenic loci QFhb-4AL and QFhb-5DL for *Fusarium* head blight resistance in common wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 20:29. doi: 10.1186/s12870-019-2177-0
- Jia, M., Yang, L., Zhang, W., Rosewarne, G., Li, J., Yang, E., et al. (2020). Genome-wide association analysis of stripe rust resistance in modern Chinese wheat. *BMC Plant Biol.* 20:491. doi: 10.1186/s12870-020-02693-w
- Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., et al. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biol.* 19:168. doi: 10.1186/s12870-019-1781-3
- Li, H. B., Xie, G. Q., Ma, J., Liu, G. R., Wen, S. M., Ban, T., et al. (2010). Genetic relationships between resistances to *Fusarium* head blight and crown rot in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 121, 941–950. doi: 10.1007/s00122-010-1363-0
- Li, X., Zhang, J. B., Song, B., Li, H. P., Xu, H. Q., Qu, B., et al. (2010). Resistance to *Fusarium* head blight and seedling blight in wheat is associated with activation of a cytochrome P450 gene. *Phytopathology* 100, 183–191. doi: 10.1094/PHYTO-100-2-0183
- Liu, P. L., Du, L., Huang, Y., Gao, S. M., and Meng, Y. (2017). Origin and diversification of leucine-rich repeat receptor-like protein kinase

- (*LRR-RLK*) genes in plants. *BMC Evol. Biol.* 17:47. doi: 10.1186/s12862-017-0891-5
- Liu, Y., Tong, H., Zhu, Z., Chen, L., Zou, J., Zhang, Y., et al. (2016). Progress in research on mechanism of resistance to fusarium head blight in wheat. *Sci. Agric. Sin.* 49, 1476–1488. doi: 10.3864/j.issn.0578-1752.2016.08.005
- Liu, Z. W., Li, H. P., Cheng, W., Peng, Y., Zhang, J. B., Gong, A. D., et al. (2012). Enhanced overall resistance to Fusarium seedling blight and Fusarium head blight in transgenic wheat by co-expression of anti-fungal peptides. *Eur. J. Plant Pathol.* 134, 721–732. doi: 10.1007/s10658-012-0048-8
- Ma, J., Stiller, J., Zhao, Q., Jian, M., Stiller, J., Qiang, Z., et al. (2014). Transcriptome and allele specificity associated with a 3bl locus for fusarium crown rot resistance in bread wheat. *PLoS One* 9:e113309. doi: 10.1371/journal.pone.0113309
- Ma, Z., Xie, Q., Li, G., Jia, H., Zhou, J., Kong, Z., et al. (2020). Germplasms, genetics and genomics for better control of disastrous wheat Fusarium head blight. *Theor. Appl. Genet.* 133, 1541–1568. doi: 10.1007/s00122-019-03525-8
- Mackey, D., Holt, B. F., Wiig, A., and Dangl, J. L. (2002). RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* 108, 743–754. doi: 10.1016/S0092-8674(02)00661-X
- Mesterhazy, A. (1987). Selection of head blight resistant wheats through improved seedling resistance. *Plant Breed.* 98, 25–36. doi: 10.1111/j.1439-0523.1987.tb01086.x
- Miedaner, T. (1997). Breeding wheat and rye for resistance to Fusarium diseases. *Plant Breed.* 116, 201–220. doi: 10.1111/j.1439-0523.1997.tb00985.x
- Nie, Y. B., and Ji, W. Q. (2019). Cloning and characterization of disease resistance protein RPM1 genes against powdery mildew in wheat line N9134. *Cereal. Res. Commun.* 47, 473–483. doi: 10.1556/0806.47.2019.27
- Ning, X., Zhang, G., Liu, X. Y., Deng, L., Cai, G. L., Zhang, Y., et al. (2010). Characterization of a novel wheat NAC transcription factor gene involved in defense response against stripe rust pathogen infection and abiotic stresses. *Mol. Biol. Rep.* 37, 3703–3712. doi: 10.1007/s11033-010-0023-4
- Perochon, A., Kahla, A., Vranić, M., Jia, J., Malla, K. B., Craze, M., et al. (2019). A wheat NAC interacts with an orphan protein and enhances resistance to Fusarium head blight disease. *Plant Biotechnol. J.* 17, 1892–1904. doi: 10.1111/pbi.13105
- Pestka, J. J., and Smolinski, A. T. (2005). Deoxynivalenol: toxicology and potential effects on humans. *J. Toxicol. Env. Health, Part B.* 8, 39–69. doi: 10.1080/10937400590889458
- Powell, J. J., Carere, J., Fitzgerald, T. L., Stiller, J., Covarelli, L., Xu, Q., et al. (2017). The Fusarium crown rot pathogen *Fusarium pseudograminearum* triggers a suite of transcriptional and metabolic changes in bread wheat (*Triticum aestivum* L.). *Ann. Bot.* 119, 853–867. doi: 10.1093/aob/mcw207
- Ren, R., Foulkes, J., Mayes, S., Yang, X., and Ray, R. V. (2016). Identification of novel quantitative trait loci for resistance to Fusarium seedling blight caused by *Microdochium majus* and *M. nivale* in wheat. *Field Crop Res.* 191, 1–12. doi: 10.1016/j.fcr.2016.03.011
- Ren, R., Yang, X., and Ray, R. V. (2015). Comparative aggressiveness of *Microdochium nivale* and *M. majus* and evaluation of screening methods for Fusarium seedling blight resistance in wheat cultivars. *Eur. J. Plant Pathol.* 141, 281–294. doi: 10.1007/s10658-014-0541-3
- Ruckenbauer, P., Buerstmayr, H., and Lemmens, M. (2001). Present strategies in resistance breeding against scab (*Fusarium* spp.). *Euphytica* 119, 123–129. doi: 10.1023/A:1017598523085
- Sapkota, S., Hao, Y., Johnson, J., Buck, J., Aoun, M., and Mergoum, M. (2019). Genome-wide association study of a worldwide collection of wheat genotypes reveals novel quantitative trait loci for leaf rust resistance. *Plant Genome* 12:190033. doi: 10.3835/plantgenome2019.05.0033
- Schilling, S., Kennedy, A., Pan, S. R., Jermini, 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.* 225, 511–529. doi: 10.1111/585232
- Shi, X. L., and Ling, H. Q. (2018). Current advances in genome sequencing of common wheat and its ancestral species. *Crop J.* 6, 15–21. doi: 10.1016/j.cj.2017.11.001
- Shin, S., Kim, K. H., Kang, C. S., Cho, K. M., Park, C. S., Okagaki, R., et al. (2014). A simple method for the assessment of Fusarium head blight resistance in Korean wheat seedlings inoculated with *Fusarium graminearum*. *Plant Pathol. J.* 30, 25–32. doi: 10.5423/PPJ.OA.06.2013.0059
- Somers, D. J., Fedak, G., and Savard, M. (2003). Molecular mapping of novel genes controlling Fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. *Genome* 46, 555–564. doi: 10.1139/g03-033
- Sorahinobar, M., Niknam, V., Ebrahimzadeh, H., Soltanloo, H., Moradi, B., and Bahram, M. (2016). Lack of association between Fusarium graminearum resistance in spike and crude extract tolerance in seedling of wheat. *Eur. J. Plant Pathol.* 144, 525–538. doi: 10.1007/s10658-015-0792-7
- Srinivasachary, Gosman, N., Steed, A., Simmonds, J., Leverington-Waite, M., Wang, Y., et al. (2008). Susceptibility to Fusarium head blight is associated with the Rht-D1b semi-dwarfing allele in wheat. *Theor. Appl. Genet.* 116, 1145–1153. doi: 10.1007/s00122-008-0742-2
- Su, J., Spears, B. J., Kim, S. H., and Gassmann, W. (2017). Constant vigilance: plant functions guarded by resistance proteins. *Plant J.* 93, 637–650. doi: 10.1111/tpj.13798
- Tamburic-Illincic, L., Somers, D., Fedak, G., and Schaafsma, A. (2009). Different quantitative trait loci for Fusarium resistance in wheat seedlings and adult stage in the Wuhan/Nyubai wheat population. *Euphytica* 165, 453–458. doi: 10.1007/s10681-008-9747-9
- Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183
- Wang, Y., and Bouwmeester, K. (2017). L-type lectin receptor kinases: new forces in plant immunity. *PLoS Pathog.* 13:e1006433. doi: 10.1371/journal.ppat.1006433
- Wang, Z., Cheng, J., Fan, A., and Zhao, J. (2018). LecRK-V, an L-type lectin receptor kinase in *Haynaldia villosa*, plays positive role in resistance to wheat powdery mildew. *Plant Biotechnol. J.* 16, 50–62. doi: 10.1111/pbi.12748
- Wiese, M. V. (1987). *Compendium of Wheat Diseases*, 2nd Edn. Saint Paul, MN: The American Phytopathological Society.
- Wisniewska, H., and Busko, M. (2005). Evaluation of spring wheat resistance to Fusarium seedling and head blight. *Biologia* 60, 287–293.
- Wu, A. B., Li, H. P., Zhao, C. S., and Liao, Y. C. (2005). Comparative pathogenicity of fusarium graminearum isolates from China revealed by wheat coleoptile and floret inoculations. *Mycopathologia* 160, 75–83. doi: 10.1007/s11046-005-1153-4
- Zhu, Z. W., Chen, L., Zhang, W., Yang, L. J., Zhu, W. W., Li, J. H., et al. (2020). Genome-Wide association analysis of fusarium head blight resistance in Chinese elite wheat lines. *Front. Plant Sci.* 11:206. doi: 10.3389/fpls.2020.00206

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 Liu, Zhu, Zhu, Chen, Niu, He, Tong, Song, Zhang, Ma and Gao. 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 Two Major QTLs in *Brassica napus* Lines With Introgressed Clubroot Resistance From Turnip Cultivar ECD01

Fengqun Yu*, Yan Zhang, Jinghe Wang, Qilin Chen, Md. Masud Karim, Bruce D. Gossen and Gary Peng

Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada, Saskatoon, SK, Canada

OPEN ACCESS

Edited by:

Harsh Raman,
New South Wales Department
of Primary Industries, Australia

Reviewed by:

Chuchuan Fan,
Huazhong Agricultural University,
China
Kun Lu,
Southwest University, China

*Correspondence:

Fengqun Yu
fengqun.yu@agr.gc.ca

Specialty section:

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

Received: 29 September 2021

Accepted: 02 December 2021

Published: 12 January 2022

Citation:

Yu F, Zhang Y, Wang J, Chen Q,
Karim MM, Gossen BD and Peng G
(2022) Identification of Two Major
QTLs in *Brassica napus* Lines With
Introgressed Clubroot Resistance
From Turnip Cultivar ECD01.
Front. Plant Sci. 12:785989.
doi: 10.3389/fpls.2021.785989

Plasmodiophora brassicae causes clubroot disease in brassica crops worldwide. *Brassica rapa*, a progenitor of *Brassica napus* (canola), possesses important sources for resistance to clubroot. A doubled haploid (DH) population consisting of 84 DH lines were developed from a Backcross2 (BC₂) plant through an interspecific cross of *B. rapa* turnip cv. ECD01 (resistant, R) with canola line DH16516 (susceptible, S) and then backcrossed with DH16516 as the recurrent parent. The DH lines and their parental lines were tested for resistance to four major pathotypes (3A, 3D, 3H, and 5X) of *P. brassicae* identified from canola. The R:S segregation ratio for pathotype 3A was 1:3, and 3:1 for pathotypes 3D, 3H, and 5X. From genotyping by sequencing (GBS), a total of 355.3 M short reads were obtained from the 84 DH lines, ranging from 0.81 to 11.67 M sequences per line. The short reads were aligned into the A-genome of *B. napus* "Darmor-bzh" version 4.1 with a total of 260 non-redundant single-nucleotide polymorphism (SNP) sites. Two quantitative trait loci (QTLs), *Rcr10*^{ECD01} and *Rcr9*^{ECD01}, were detected for the pathotypes in chromosomes A03 and A08, respectively. *Rcr10*^{ECD01} and *Rcr9*^{ECD01} were responsible for resistance to 3A, 3D, and 3H, while only one QTL, *Rcr9*^{ECD01}, was responsible for resistance to pathotype 5X. The logarithm of the odds (LOD) values, phenotypic variation explained (PVE), additive (Add) values, and confidence interval (CI) from the estimated QTL position varied with QTL, with a range of 5.2–12.2 for LOD, 16.2–43.3% for PVE, 14.3–25.4 for Add, and 1.5–12.0 cM for CI. The presence of the QTLs on the chromosomes was confirmed through the identification of the percentage of polymorphic variants using bulked-segregant analysis. There was one gene encoding a disease resistance protein and 24 genes encoding proteins with function related to plant defense response in the *Rcr10*^{ECD01} target region. In the *Rcr9*^{ECD01} region, two genes encoded disease resistance proteins and 10 genes encoded with defense-related function. The target regions for *Rcr10*^{ECD01} and *Rcr9*^{ECD01} in *B. napus* were homologous to the 11.0–16.0 Mb interval of chromosome A03 and the 12.0–14.5 Mb interval of A08 in *B. rapa* "Chiifu" reference genome, respectively.

Keywords: *Brassica napus*, *Brassica rapa*, *Plasmodiophora brassicae*, clubroot, genotyping by sequencing, ECD01, resistance, pathotype

INTRODUCTION

Brassica species are grown for the production of edible oil and vegetables. The genomic relationships among the main species of brassica crops were explained by the “triangle of U” (Morinaga, 1934; Nagaharu and Nagaharu, 1935); *Brassica rapa* (genome represented as AA; $n = 10$), *Brassica nigra* (BB; $n = 8$), and *Brassica oleracea* (CC; $n = 9$) are diploid species, and *Brassica napus* (AACC; $n = 19$), *Brassica juncea* (AABB; $n = 18$), and *Brassica carinata* (BBCC; $n = 17$) are amphidiploid species resulting from hybridization between pairs of the diploid species.

Clubroot, caused by the obligate soil-borne pathogen *Plasmodiophora brassicae* Woronin, is an important disease in brassica crops worldwide. The pathogen belongs to the infrakingdom Rhizaria, a diverse group of amoeboid microbes (Nikolaev et al., 2004). Root infection by *P. brassicae* results in the formation of characteristic clubs, also known as “galls,” on the roots of host plants. These abnormal growths restrict the flow of water and nutrients to the plant, resulting in above-ground symptoms that include stunting, yellowing, premature senescence, and reduction in both seed yield and quality (Pageau et al., 2006). *B. napus* (oilseed rape/canola) is an important crop for edible oil production worldwide. Clubroot was first identified in canola fields on the Canadian Prairies in 2003 but has spread rapidly to pose a serious threat to canola production in Canada.

Strains of *P. brassicae* collected in Canada have been classified into more than 30 pathotypes based on the reactions on the Canadian Clubroot Differential (CCD) set (Strelkov et al., 2018; Hollman et al., 2021). Among the pathotypes, 3H was the most prevalent original pathotype, 5X was the first new pathotype that was aggressive on the first generation of Canadian clubroot-resistant cultivars, and 3A and 3D are currently the predominate new pathotypes in Alberta (Dakouri et al., 2021). The pathogen can survive in soil as resting spores for a long period, so it is difficult to manage using cultural practices or chemical treatments (Voorrips, 1995). Genetic resistance can be an effective strategy for clubroot management, but the sources available for resistance to clubroot in *B. napus* are very limited. Strong resistance was identified in its progenitor species, *B. rapa*, especially in European turnip, *B. rapa* subsp. *rapifera*, which was reviewed by Hirai (2006). The resistance to clubroot available from European turnips has been transferred into Chinese cabbage (*B. rapa*) (Piao et al., 2009). Introgression of traits from turnip into *B. napus* is possible via interspecific crosses, so turnip has been a valuable source for resistance to clubroot in canola. Clubroot resistance (CR) from turnip cultivar “Debra” has been transferred into *B. napus* cultivars of swede (Lammerink, 1970) and from turnip cultivar “Waaslander” (also known as ECD04) into forage (Johnston, 1974) and oilseed lines of *B. napus* (Gowers, 1982).

Genetic mapping of CR genes is an important step toward breeding for resistance to clubroot. To date, more than 20 genes or quantitative trait locus (QTLs) have been mapped to six chromosomes of the A-genome in *B. rapa* through biparental mapping methods. *Crr2* and *PbBa1.1* were located on A01 (Suwabe et al., 2003; Chen et al., 2013); *CRc* and *Rcr8* were located on A02 (Sakamoto et al., 2008; Yu et al., 2017);

Bra.CR.a, *Bra.CR.c*, *Crr3*, *CRA*, *CRb*, *CRb^{kato}*, *CRd*, *CRk*, *PbBa3.1*, *PbBa3.2*, *PbBa3.3*, *Rcr1*, *Rcr2*, *Rcr4*, and *Rcr5* were located on A03 (Matsumoto et al., 1998; Hirai et al., 2004; Piao et al., 2004; Sakamoto et al., 2008; Chen et al., 2013; Chu et al., 2014; Pang et al., 2014; Yu et al., 2016, 2017; Huang et al., 2017, 2019; Hirani et al., 2018); *Crr4* were located on A06 (Suwabe et al., 2003); *qBrCR38-1* were located on A07 (Zhu et al., 2019); *Crr1*, *CRs*, *PbBa8.1*, *Bra.CR.b*, *Rcr3*, *Rcr9/Rcr9^{wa}*, and *qBrCR38-2* were located on A08 (Suwabe et al., 2006; Chen et al., 2013; Yu et al., 2017; Hirani et al., 2018; Laila et al., 2019; Zhu et al., 2019; Karim et al., 2020). Three genes, *Crr1*, *CRA*, and *CRb^{kato}*, have been cloned, all of which encode toll-interleukin-1 receptor, nucleotide binding site, and leucine-rich repeat (TIR-NBS-LRR, TNL) proteins (Ueno et al., 2012; Hatakeyama et al., 2013, 2017). The identification of CR genes has been also carried out in *B. oleracea* (Lee et al., 2016; Dakouri et al., 2018; Peng et al., 2018), *B. nigra* (Chang et al., 2019), and *B. napus* (Manzanares-Dauleux et al., 2000; Werner et al., 2008; Fredua-Agyeman and Rahman, 2016; Hasan and Rahman, 2016; Botero-Ramírez et al., 2020).

Brassica rapa turnip cv. “Debra” was used as the donor for developing CR in swede cultivars (Lammerink, 1970). “Debra” was included in the European clubroot differential (ECD) set as differential line ECD01 (Buczacki et al., 1975; Diederichsen et al., 2009). *CRb*, a CR gene identified in a Chinese cabbage cv. “CR Shinki” and two CR genes in Chinese cabbage cv. “CR Kanko,” *CRk* and *CRc*, were derived from ECD01 (Piao et al., 2004) and “Debra” (Sakamoto et al., 2008), respectively. Two other CR genes, *BraA.CR.a* (A03) and *BraA.CR.b* (A08), were also identified from ECD01 (Hirani et al., 2018). Finally, ECD01 was resistant to all of the Canadian pathotypes of *P. brassicae* described by Strelkov et al. (2018) (Yu F, unpublished data), which makes it a valuable source of genes for CR canola in Canada.

In this study, an interspecific cross of ECD01 \times *B. napus* line DH16516 was made, and the resulting F₁ progeny were backcrossed with DH16516 to produce BC₁. Continuing backcross was made by crossing the BC₁ with DH16516 to produce Backcross2 (BC₂). A doubled haploid (DH) population consisting of 84 DH lines from a single BC₂ plant was developed. Genotyping by sequencing (GBS) analysis of the A-genome of *B. napus* was used to (1) characterize the genome-wide DNA variants in the DH lines, (2) detect QTLs associated with resistance to the most important pathotypes of *P. brassicae* on the Canadian Prairies, and (3) identify putative candidate genes for each QTL.

MATERIALS AND METHODS

Plant Materials

A seed of ECD01, a turnip (*B. rapa*) cultivar carrying genes for CR, was provided by Nutrien Ag Solutions (Saskatoon, SK, Canada). DH16516 is a spring-type, clubroot-susceptible, DH canola-quality line of *B. napus* developed by Dr. Séguin-Swartz at Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada (SRDC, AAFC), Saskatoon, SK, Canada. ECD01 was crossed to DH16516 (pollen donor) to produce F₁

progeny. Backcrosses with DH16516 (recurrent parent) were performed to produce the BC₁ and BC₂ populations. A BC₂ plant with resistance to pathotype 5X was chosen as the donor for microspore culture. In total, 84 DH lines were developed by HaploTech Inc. (Winnipeg, Canada) through a fee-for-service contract. Seed from three plants of each DH line was increased in a greenhouse at SRDC for study.

Evaluation of Resistance to Clubroot

Clubroot strains collected in canola fields in Alberta were characterized based on pathotype and provided by Dr. S. E. Strelkov at the University of Alberta, Canada. The method and experimental design used in this study were as described by Suwabe et al. (2003). Plants were tested for resistance to four pathotypes of *P. brassicae* (strain F.3-14 for pathotype 3A, F.1-14 for 3D, P. 41-14 for 3H, and LG02 for 5X). Fresh and clean clubbed roots harvested at 4–5 weeks after inoculation of each strain were cut into smaller pieces with scissors, macerated in distilled water for 1–2 h, and blended in a blender at high speed for 2 min. After filtering through eight layers of sterile cheesecloth, resting spores extracted from the clubbed roots were adjusted to a concentration of 1.0×10^7 resting spores/ml in distilled water for plant inoculation.

Seeds of the DH population were sown into Sunshine #3 soilless mix (Sun Gro Horticulture Canada Ltd., Seba Beach, AB, Canada) with Osmocote (Everris NA Inc., Dublin, OH, United States) in 32 pot inserts held by trays (The HC Companies, Twinsburg, OH, United States). Approximately 4 L of water was added to each tray to soak the soilless mix overnight. Seven days after planting, inoculation was performed by adding 15 ml of inoculum (1×10^7 spores/ml) into each pot with 6–9 seedlings of each line. The inoculated plants were grown in a growth chamber set at 22/18°C day/night temperature with a 16-h photoperiod. The canola cultivar “45H29” (resistant to pathotype 3H) and the parental lines (ECD01 and DH16516) were included as controls. Six weeks after inoculation, plants were pulled and the roots were examined for clubroot symptoms.

Clubroot severity was evaluated on a 0–3 scale, where 0 indicates no clubbing, 1 indicates a few small clubs, 2 indicates moderate clubbing, and 3 indicates severe clubbing. A disease severity index (DSI) was calculated for each line using the method of Horiuchi and Hori (1980):

$$DSI = \frac{\sum (\text{rating class}) \times (\# \text{ plants in rating class})}{\text{total } \# \text{ plants in treatment} \times 3} \times 100$$

Correlation coefficients of severity among the DH families to four pathotypes of *P. brassicae* were calculated in Microsoft Excel function “Correl” using the equation:

$$\text{Correl}(X, Y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Significance was determined using *t*-tests (Iversen and Gergen, 1997). Each line with a resistance response ($DSI \leq 30\%$) in the initial study was reassessed two more times. Each of these repetitions provided a similar result in most cases. For those

lines with inconsistent results, the highest DSI among the three repetitions of the assessment was considered to be the most accurate and was used to characterize the resistance response of the line. DH lines with $DSI \leq 30\%$ were classified as R and those lines with $DSI > 30\%$ as S lines.

The F₁ plants were tested with pathotype 3H, and the BC₂ donor plant was only tested with 5X for several reasons. First, the clubroot reaction of a single plant can be assessed for only one pathotype. Second, pathotype 3H was the predominant pathotype in Canada when we obtained the F₁ progeny and performed the selection for CR. Similarly, at that time when the donor plant from BC₂ was chosen for microspore culture, pathotype 5X was the only new pathotype that had been identified. Third, only a few seeds of F₁ were obtained due to difficulties in the interspecific cross, so it was not possible to test multiple pathotypes in the F₁ progeny.

DNA Sequencing and Alignment of Reads to a Reference Genome

DNA was extracted from young leaves of each of the 84 DH lines and parental lines following the DNeasy Plant Mini Handbook from QIAGEN. GBS of the 84 DNA samples and two replications of the parental cultivar ECD01 were performed on an Illumina platform with pair-end sequencing at BGI Americas Corp (Cambridge, MA, United States). Two replications of cv. ECD01 were performed to increase the sequencing depth for this parental line to provide a more accurate call of the genotype at each single-nucleotide polymorphism (SNP) site in the DH population. DH16516 is an important *B. napus* canola recipient line for introgression of CR at AAFC, Saskatoon, so whole-genome sequencing of the line had already been performed at Plant Biotechnology Centre (Saskatoon, SK, Canada) as part of the generation of a new reference genome (unpublished data). The short reads from the whole-genome sequencing data were used for this study. The program SeqMan NGen 15 (DNASTAR, Madison, WI, United States) was used for short read assembly. “Whole genome DNA-Seq/Genotyping” assembly workflow, “Reference based assembly-normal workflows” assembly type, and “Automatic Mer size, Automatic Minimum match percentage, High Layout stringency and Medium SNP filtering stringency” assembly options were chosen. Short reads from each of the 84 DH samples, parental DH16516, and the combined two replicates of ECD01 were aligned to *B. napus* reference genome for cv. “Darmor-bzh” version 4¹.

Identification of Variants, Variant Filtering, Construction of Linkage Map, and Quantitative Trait Locus Mapping

Identification of variants (SNPs and InDels) in the DNA sequences of each BC₂ DH sample relative to the reference genome of *B. napus* “Darmor-bzh” was performed using SeqMan Pro 15 (DNASTAR, Madison, WI, United States), but only SNPs were used for further study. Comparison of the

¹<https://www.genoscope.cns.fr/brassicnapus/data/>

variants among the 84 BC₂ DH samples was carried out using Qseq 15 (DNASTAR).

Genotyping by sequencing-SNP sites were named based on the reference genome (DM: “Darmor-*bzh*”), the A-genome chromosome (A01–A10), and the position on the reference chromosome sequence. An SNP site was called in a given sample at following criteria: depth > 5, $Q > 30$, and SNP percentage > 50%. Since the recipient parent DH16516 and the 84 DH lines were DH lines, all SNP sites should theoretically be homozygous. After filtering, heterozygous genotypes in the parental line DH16516 and the DH lines and monomorphic phenotypes between the parents or among the 84 individuals were removed.

The remaining SNP sites after filtering were further analyzed using JoinMap 4.1 (Kyazma B.V., Lelystad, The Netherlands; Van Ooijen, 2011). SNP alleles from the resistant parent (ECD01) were scored as “B,” and those from the susceptible parent (DH16516) as “A.” Marker orders and positions in the genetic map were determined using maximum likelihood in the Kosambi’s model with a minimum logarithm of the odds (LOD) values of 10. Only SNP sites that could be assigned into the 10 chromosomes of the A-genome at LOD scores of 10.0 were kept. The set of filtered SNP sites obtained was used for binning of redundant markers, construction of linkage map, and mapping of QTLs for resistance to clubroot using the QTL IciMapping Inclusive Composite Interval Mapping (ICIM) method (Meng et al., 2015). A linkage map was drawn using MapChart 2.1 (Droevendaalsesteeg 4, Wageningen, Netherlands; Voorrips, 2002) based on the genetic location determined with QTL IciMapping. The LOD score threshold was set using a 1,000-permutation test with a type I error of 0.05 for QTL declaration. The QTL effects were estimated as phenotypic variation explained (PVE) and additive (Add) values by each QTL.

Identification of Genes in the Target Regions of the *B. napus* “Darmor-*bzh*” Reference Genome

Gene annotation was analyzed using Blast2GO (Conesa et al., 2005) using coding sequences (CDS) of the genes in each of the QTL target regions from 1 Mb upstream to 1 Mb downstream of the SNP markers in the peak regions as determined by IciMapping. Genes related to disease resistance and defense responses were identified using Blast2GO information of the gene description and gene ontology. The most probable Arabidopsis homolog corresponding to each disease resistance gene and the class of disease resistance proteins were obtained using the CDS of the disease resistance gene in the *B. napus* by Blast search at www.arabidopsis.org.

Mapping of the Quantitative Trait Loci With Bulk Segregant Analysis

Bulked segregant analysis (BSA) has been used to detect molecular markers linked to traits of interest, such as disease resistance (Michelmore et al., 1991). In BSA, bulks of plants with contrasting phenotypes are generated. Our previous studies showed that a gene could be genetically mapped by identifying

the percentage of polymorphic variants (PPV) in a genome using BSA (Yu et al., 2016; Dakouri et al., 2018; Huang et al., 2019; Karim et al., 2020).

Doubled haploid lines were selected to form a R bulk and a S bulk based on their phenotypes using SNP marker-assisted selection. GBS data from the R and S bulks were aligned onto the *B. napus* reference genome separately using SeqMan NGen 15 (DNASTAR). Mapping of the QTLs was performed using the PPV method described by Yu et al. (2016) and Dakouri et al. (2018).

Search for the Syntenic Regions of Identified Quantitative Trait Loci in *B. rapa* “Chiifu” Reference Genome

The *B. rapa* reference genome version 3.0 (Zhang et al., 2018) was downloaded from <https://brassicadb.org/brad/downloadOverview.php>. DNA sequences of the QTL target regions from the A-genome of *B. napus* were aligned into the *B. rapa* genome using MegAlign Pro 15 with MAUVE (DNASTAR).

RESULTS

Resistance to Clubroot in the Parental Lines and the Backcross2 Doubled Haploid Population

The clubroot reaction of the parental lines (ECD01 and DH16516), controls, and the DH population was assessed against pathotypes 3A, 3D, 3H, and 5X (Table 1). As expected, ECD01 was highly resistant to all pathotypes (0% DSI), DH16516 was highly susceptible (100% DSI), and “45H29” was resistant to pathotype 3H only (Figure 1 and Table 1). The F₁ plants from the interspecific crosses of DH16516 × ECD01 were highly resistant to pathotype 3H (0% DSI), which was the predominate pathotype in Canada before the emergence of the 3A, 3D, and 5X. Clubroot severity in response to inoculation with each pathotype in the DH population could be divided into two classes: resistant (R) lines with DSI ≤ 30% and susceptible (S) lines with DSI > 30% (Figure 2). The segregation ratio of R and S was calculated, and the goodness-of-fit was tested with a χ^2 test using Microsoft Excel software. Of the four pathotypes, segregation of R and S best fit

TABLE 1 | Genetic analysis of resistance of the parental lines (DH16516, ECD01), controls (cv. “45H29”), and the inoculation of doubled haploid (DH) population derived from BC₂ with four pathotypes of *Plasmodiophora brassicae* based on the clubroot severity (disease severity index, DSI) of each line (Resistant, R, DSI ≤ 30; Susceptible, S, DSI > 30).

| Patho type | DSIs | | | | No. of DH lines | | | P-value of ratio | | |
|------------|-------|---------|----------------|-------|-----------------|----|----|------------------|-------|-------|
| | ECD01 | DH16516 | F ₁ | 45H29 | Total | R | S | 1:1 | 3:1 | 1:3 |
| 3A | 0 | 100 | – | 100 | 82 | 27 | 55 | 0.001 | 0.001 | 0.100 |
| 3D | 0 | 100 | – | 100 | 80 | 49 | 31 | 0.001 | 0.005 | 0.001 |
| 3H | 0 | 100 | 0 | 0 | 82 | 61 | 21 | 0.001 | 0.90 | 0.001 |
| 5X | 0 | 100 | – | 100 | 84 | 52 | 32 | 0.001 | 0.006 | 0.001 |

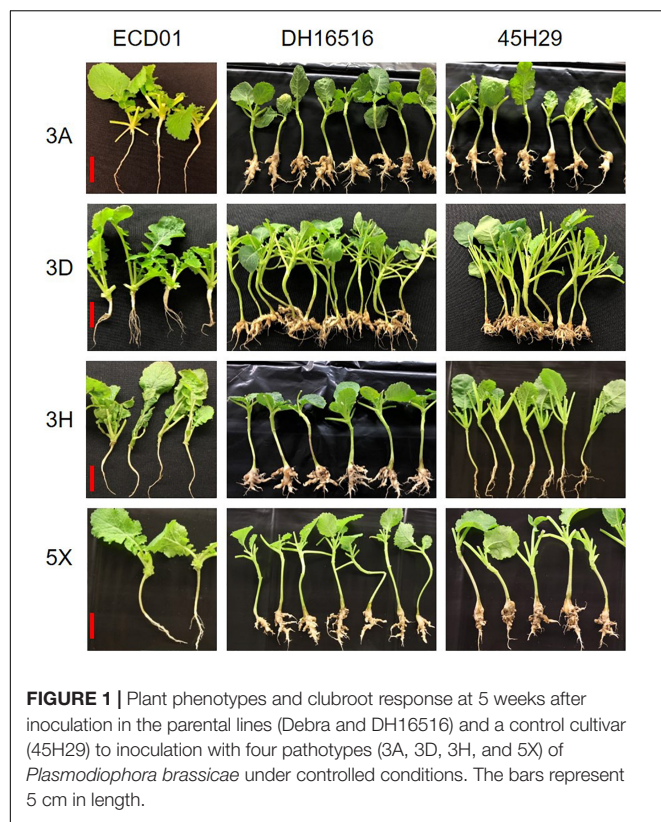


FIGURE 1 | Plant phenotypes and clubroot response at 5 weeks after inoculation in the parental lines (Debra and DH16516) and a control cultivar (45H29) to inoculation with four pathotypes (3A, 3D, 3H, and 5X) of *Plasmodiophora brassicae* under controlled conditions. The bars represent 5 cm in length.

a 1:3 ratio for pathotype 3A and a 3:1 ratio for pathotypes 3D, 3H, and 5X. These results indicated that resistance to pathotype 3A was controlled by two genes in complementary action, and resistance to pathotypes 3D, 3H, and 5X was controlled by two genes in duplicate action.

Correlation coefficients among the DSI values for the pathotypes ranged from 0.55 to 0.81, but all were significant at $P < 0.01$ (Table 2). This indicated that the genes for resistance to the different pathotypes were likely controlled by the same genes or tightly linked genes.

Alignment of DNA Short Reads Into the *B. napus* Genome

Since CR in the DH population originated from the A-genome of *B. rapa* cv. ECD01, only A-genome DNA sequences in the reference genome *B. napus* “Darmor” version 4.1 were used for alignment of DNA short reads and discovery of DNA variants (SNPs and InDels). Approximately 219.9 million (M) short reads were obtained from whole-genome sequencing from DH16516, and 53.1% of the reads were assembled into the reference A-genome; 13.8 M sequences were obtained from GBS of ECD01, and 70.9% were assembled into the reference A-genome. A total of 355.3 M short reads from 84 DH lines were obtained, ranging from 0.81 to 11.67 M sequences per line (Supplementary Figure 1). The mean number of reads aligned into the reference genome from each line was 2.3 M (range 0.46–5.22 M, Supplementary Figure 1), and 54.7% were assembled into the reference A-genome.

Identification of Polymorphic Single-Nucleotide Polymorphism Sites and Quantitative Trait Locus Analysis

After the initial filtering, 429 polymorphic SNP sites were left and were distributed to 9 of 10 chromosomes of the reference genome of “Darmor-bzh” (Supplementary Table 1). No polymorphic markers were identified from chromosome A06. There was no correlation between chromosome size and the number of SNP markers identified ($r = -0.092$) in the population. To remove redundant markers, the 429 SNP sites were further filtered using the binning function in IciMapping, which left only 260 non-redundant SNP sites (Table 3). A genetic map of the nine chromosomes of the A-genome was constructed from the distributed SNP sites (Supplementary Figure 2). The length of each chromosome ranged from 0 (chromosome A06) to 471.8 cM (A01), with an average length of 85.3 cM. Chromosome A01 was much longer than the other linkage groups. The number of SNP sites per chromosome ranged from 0 (A06) to 152 (A01), with a mean of 26 SNPs per chromosome. The SNP interval of each chromosome ranged from 0.8 to 4.8 cM, with a mean of 3.3 cM (Supplementary Table 1).

Mapping of the QTLs was performed using the linkage map (Supplementary Figure 1) and trait values for resistance to each pathotype (3A, 3D, 3H, and 5X). Two QTLs were identified: a QTL designated as *Rcr10*^{ECD01} on A03, with a peak at the SNP markers DM_A03_12570715 and DM_A03_10873502, and a QTL designated as *Rcr9*^{ECD01} (Figure 3), located near the previously identified genes *Rcr9* and *Rcr9*^{wa} (Yu et al., 2017; Karim et al., 2020) on A08, with a peak at DM_A08_10325589 and DM_A08_10529713 (Table 3). Resistance to pathotypes 3A, 3D, and 3H was associated with the two QTLs (*Rcr10*^{ECD01} and *Rcr9*^{ECD01}), but resistance to 5X was only associated with *Rcr9*^{ECD01} (Table 3). LOD, PVE, Add values, and CI from the estimated QTL position varied between the QTLs, ranging from 5.2 to 12.2 for LOD, 16.2 to 43.3% for PVE, 14.6 to 25.4 for Add, and 1.5 to 12 cM for CI (Table 3). The values of Add for the two QTLs were positive, indicating that the resistant loci were derived from the resistant parent ECD01.

Identification of Disease Resistance Genes and Genes Related to Plant Defense Response

Searches for candidate genes for *Rcr10*^{ECD01} and *Rcr9*^{ECD01} that encoded disease resistance proteins and defense-related genes were performed using CDS of the reference genome in the target region including 1 Mb up- and downstream of the left and right markers (Table 3).

Rcr10^{ECD01}, which was responsible for resistance to pathotypes 3A, 3D, and 3H, was mapped into chromosome A03, with a peak at SNP markers DM_A03_10873502 and DM_A03_12570715 (Table 3). There are 676 genes in this 3.7 Mb region (Supplementary Table 3). Among the genes, one gene (*BnaA03g25330D*) encoded a disease resistance protein (Table 4), and 24 genes encoded proteins with functions related to plant defense response (Supplementary Table 3). *BnaA03g25330D* is

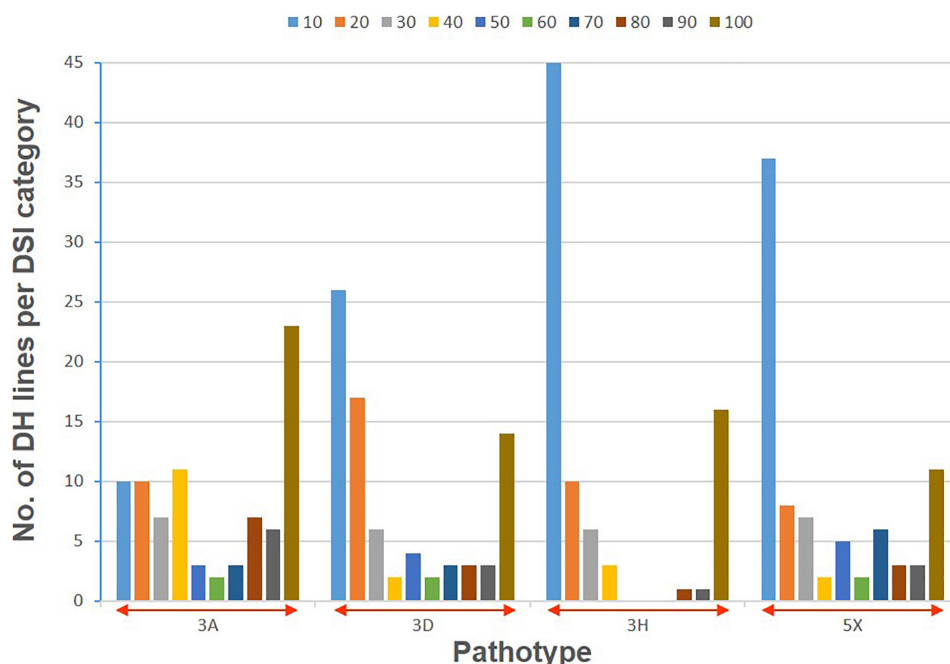


FIGURE 2 | Distribution of clubroot severity (disease severity index, DSI) following inoculation with four pathotypes (3A, 3D, 3H, and 5X) of *P. brassicae* in a doubled haploid (DH) population derived from a BC₂ plant of *Brassica rapa* ECD01 crossed with *B. napus* DH16516. Colors in each stacked column represent the proportion of the lines with a DSI value within that decile (= 10% range).

homologous to the Arabidopsis gene *AT5G22690*, which encoded a TNL protein (Table 4).

Rcr9^{ECD01}, which was responsible for resistance to all four pathotypes, was mapped into chromosome A08, with a peak at SNP markers DM_A08_10325589 and DM_A08_10529713. There were 338 genes in this 2.2 Mb region (Table 4 and Supplementary Table 3). Two genes (*BnaA08g10100D* and *BnaA08g11840D*) encoded disease resistance proteins, and *BnaA08g10100D* was homologous to the previously cloned resistance gene *Crr1*. *BnaA08g10100D* and *BnaA08g11840D* were homologous to the Arabidopsis genes *AT5G11250* and *AT4G33300*, respectively. *AT5G11250* encodes an atypical TNL protein and *AT4G33300* encodes a member of the activated disease resistance 1 family nucleotide-binding leucine-rich repeat immune receptors (Table 4). Also, this region contained 10 genes that encoded proteins with defense-related functions (Supplementary Table 3).

TABLE 2 | Correlation coefficients for clubroot severity after inoculation of DH population derived from BC₂ of DH16516 × ECD01 for resistance to four pathotypes of *P. brassicae*.

| Pathotype | 3A | 3D | 3H | 5X |
|-----------|--------|--------|--------|------|
| 3A | 1.00 | | | |
| 3D | 0.64** | 1.00 | | |
| 3H | 0.65** | 0.81** | 1.00 | |
| 5X | 0.55** | 0.55** | 0.68** | 1.00 |

**Significance level at $P < 0.01$.

Confirming the Quantitative Trait Locus Intervals With Bulk Segregant Analysis

Of the 84 DH lines, 19 lines were resistant to almost all the pathotypes. They all carried alleles from the resistant parent ECD01 (SNP genotype “B”) with *Rcr10^{ECD01}* (DM_A03_10873502 and DM_A03_12570715) and *Rcr9^{ECD01}* (DM_A08_10325589 and DM_A08_10529713). Also, 17 lines were susceptible to almost all the pathotypes and all of them carried alleles from the susceptible parent line DH16516 (SNP genotype “A”) for the two QTLs. As a result, the R bulk was formed from the 19 R DH lines, while the S bulk was formed from the 17 S DH lines for the BSA (Supplementary Table 4).

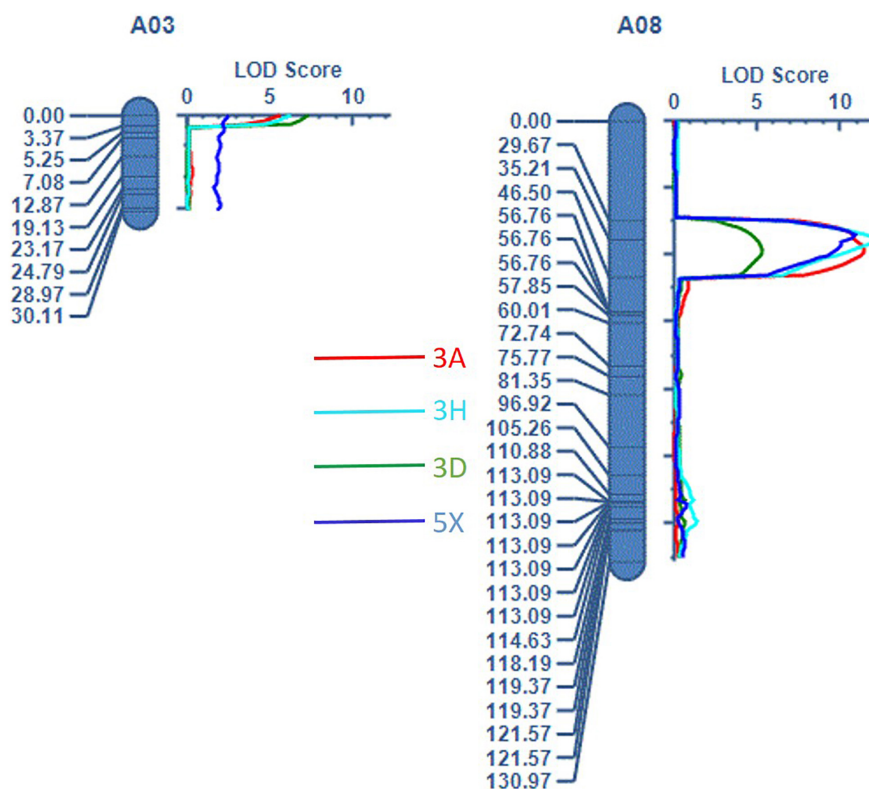
A total of 93.5 M short reads from the R bulk and 69.4 M short reads from the S bulk were aligned into the *B. napus* reference genome. A PPV peak (25–30%) occurred within the physical interval 9–14 Mb on chromosome A03 and the other peak (25–36%) within the physical interval 9–12 Mb on chromosome A08 (Figure 4), which indicated that *Rcr10^{ECD01}* and *Rcr9^{ECD01}* resided in the intervals of chromosomes A03 and A08, respectively. This result is consistent with that from the above QTL analysis.

Search for the Syntenic Regions of the Quantitative Trait Loci in the *B. rapa* “Chiifu” Reference Genome

Most of the genes or QTLs for CR in *Brassica* species containing the A-genome that have been identified were from *B. rapa*. In this study, the DH population was developed with introgression of

TABLE 3 | QTL position, phenotypic variation explained (PVE), additive (Add), the logarithm of the odds (LOD), and confidence interval (CI) for the QTLs originating from *Brassica rapa* ECD01 for resistance to four pathotypes of *P. brassicae* (permutations = 1,000).

| Pathotype | Chromosome/QTL | Position | Left marker | Right marker | LOD | PVE (%) | Add | Left CI | Right CI |
|-----------|------------------------------------|----------|-----------------|-----------------|------|---------|------|---------|----------|
| 3A | A03/ <i>Rcr10</i> ^{ECD01} | 0 | DM_A03_12570715 | DM_A03_10873502 | 5.6 | 16.2 | 14.3 | 0.0 | 1.5 |
| | A08/ <i>Rcr9</i> ^{ECD01} | 38 | DM_A08_10325589 | DM_A08_10529713 | 11.4 | 43.3 | 22.4 | 35.5 | 42.5 |
| 3D | A03/ <i>Rcr10</i> ^{ECD01} | 0 | DM_A03_12570715 | DM_A03_10873502 | 7.2 | 27.3 | 21.4 | 0.0 | 2.5 |
| | A08/ <i>Rcr9</i> ^{ECD01} | 38 | DM_A08_10325589 | DM_A08_10529713 | 5.2 | 21.5 | 18.1 | 32.5 | 44.5 |
| 3H | A03/ <i>Rcr10</i> ^{ECD01} | 0 | DM_A03_12570715 | DM_A03_10873502 | 6.2 | 17.9 | 16.6 | 0 | 2.5 |
| | A08/ <i>Rcr9</i> ^{ECD01} | 35 | DM_A08_10337601 | DM_A08_10325589 | 12.2 | 43.1 | 25.2 | 33.5 | 37.5 |
| 5X | A08/ <i>Rcr9</i> ^{ECD01} | 34 | DM_A08_10337601 | DM_A08_10325589 | 10.9 | 42.8 | 25.4 | 32.5 | 37.5 |

**FIGURE 3** | Two QTLs were detected: *Rcr10*^{ECD01} on chromosome A03 and *Rcr9*^{ECD01} on A08.

QTLs from *B. rapa*, so the QTL target regions of chromosome A03 and A08 of *B. napus* were compared with those of *B. rapa*.

The *B. rapa* reference genome “Chiifu” version 3.0 is the most recent version available for the “Chiifu” reference genome (Zhang et al., 2018). The 3.7 Mb region from 9.8 to 13.5 Mb of *B. napus* chromosome A03, which included a fragment of the markers DM_A03_10873502 and DM_A03_12570715 for *Rcr10*^{ECD01}, was homologous to the region 11.0–16 Mb of “Chiifu” A03 (Figure 5). *Rcr9*^{ECD01}, located on the 2.2 Mb length from 9.3 to 11.5 Mb of *B. napus* chromosome A08, which included SNP markers DM_A08_10325589 and DM_A08_10529713, was homologous to the region 12.0–14.5 Mb of A08 in *B. rapa* “Chiifu” (Figure 5).

DISCUSSION

Clubroot has the potential to become an important constraint to canola production on the Canadian Prairies. Pathotype 3H was (and likely still is) the predominant pathotype in the Prairie region, pathotype 5X was the first new pathotype identified as virulent on resistant canola cultivars such as “45H29,” and 3A and 3D have become the most prevalent among the new virulent pathotypes (Hollman et al., 2021). Therefore, these four pathotypes were selected for this study.

Two QTLs for resistance to the four pathotypes of *P. brassicae* derived from *B. rapa* ECD01 were transferred to, identified, and mapped in a DH population of *B. napus*. The DH population was segregated in a 1:3 (R:S) ratio for resistance to pathotype 3A. This indicated that resistance to pathotype 3A was controlled

TABLE 4 | A list of genes encoding proteins associated with plant disease resistance through BLAST2GO and Blast searches with CDS in the QTL target regions at <https://www.arabidopsis.org/Blast/index.jsp>.

| QTL | <i>Rcr10^{ECD01}</i> | <i>Rcr3/9^{ECD01}</i> | |
|--------------------------------------|--|--|--|
| R to pathotype | 3A, 3D, and 3H | 3A, 3D, 3H, and 5X | |
| Chromosome | A03 | A08 | |
| Gene name | <i>BnaA03g25330D</i> | <i>BnaA08g10100D</i> | <i>BnaA08g11840D</i> |
| <i>B. napus</i> gene location (base) | 12234711...12240552 | 9456084...9467947 | 10622229...10625339 |
| Length (base) | 5841 | 11863 | 3110 |
| Description from Blast2GO | Disease resistance protein RPS6-like | Disease resistance protein TAO1-like | Probable disease resistance protein At4g33300 |
| Function with Blast2GO | Hydrolase activity; ADP binding; defense response; signal transduction | Hydrolase activity; ADP binding; defense response; signal transduction | ADP binding |
| Homolog in Arabidopsis | <i>AT5G22690</i> | <i>AT5G11250</i> | <i>AT4G33300</i> |
| R gene class | Disease resistance protein (TIR-NBS-LRR class) family | TIR-NBS-LRR protein involved in stress response | Activated disease resistance 1 (ADR1) family of NBS-LRR immune receptors |

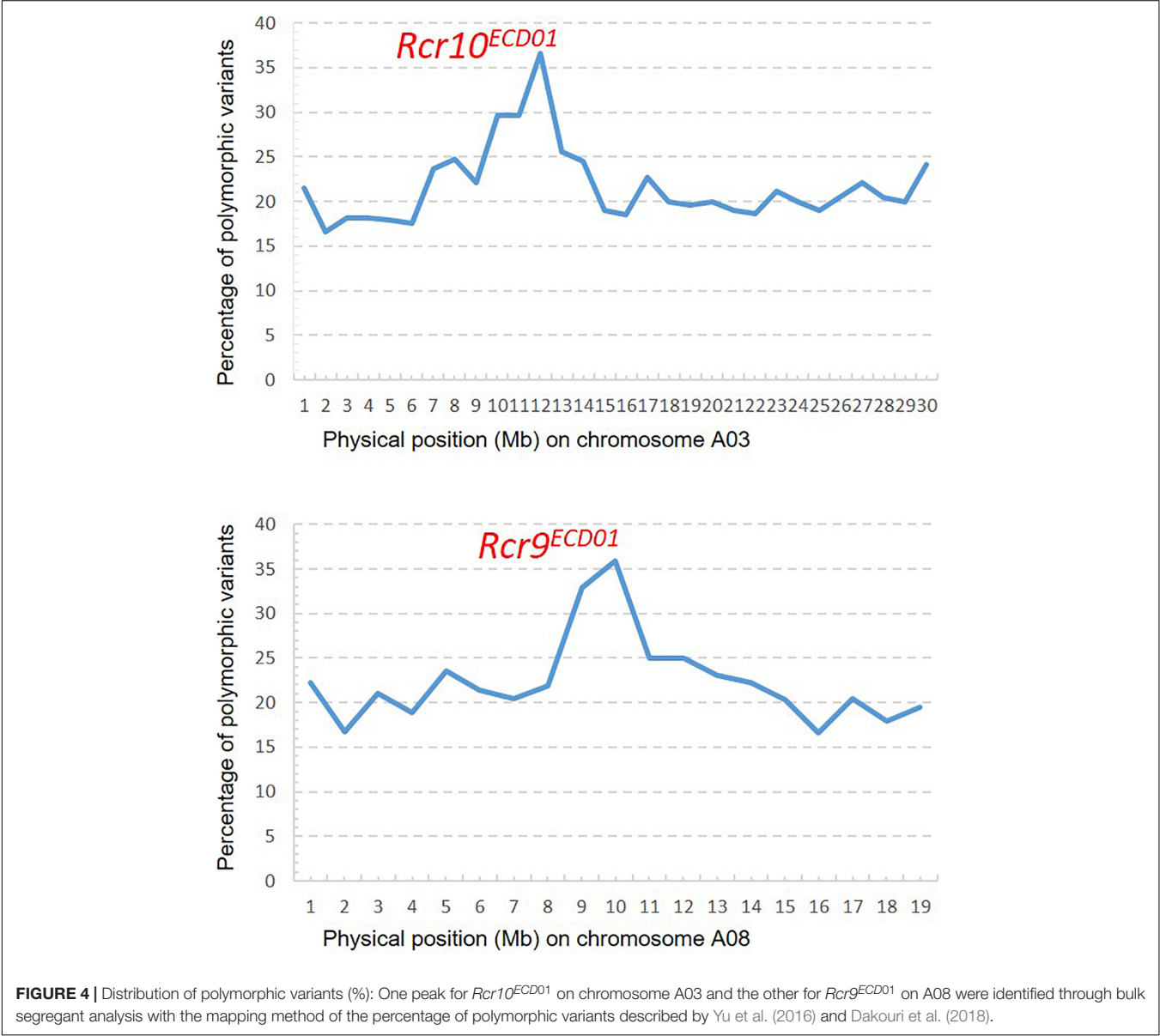


FIGURE 4 | Distribution of polymorphic variants (%): One peak for *Rcr10^{ECD01}* on chromosome A03 and the other for *Rcr9^{ECD01}* on A08 were identified through bulk segregant analysis with the mapping method of the percentage of polymorphic variants described by Yu et al. (2016) and Dakouri et al. (2018).

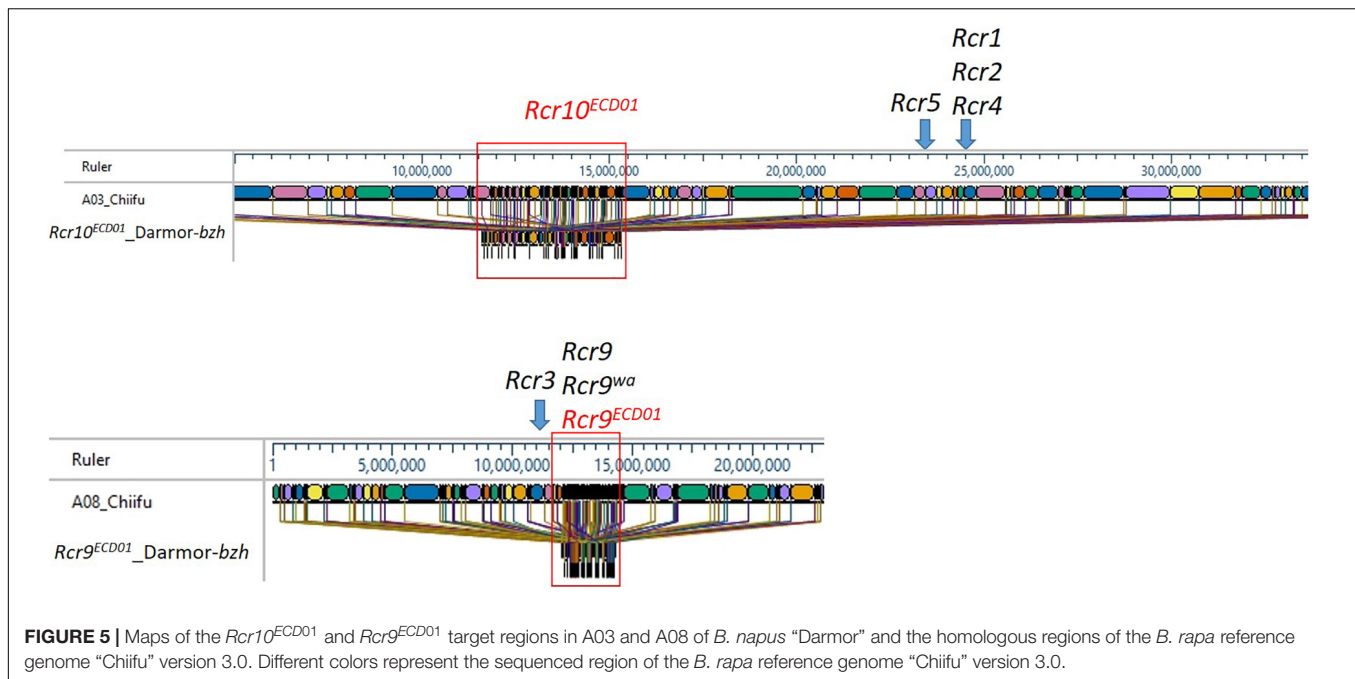


FIGURE 5 | Maps of the *Rcr10*^{ECD01} and *Rcr9*^{ECD01} target regions in A03 and A08 of *B. napus* “Darmor” and the homologous regions of the *B. rapa* reference genome “Chiifu” version 3.0. Different colors represent the sequenced region of the *B. rapa* reference genome “Chiifu” version 3.0.

by two genes in complementary action. The segregation ratio for resistance to pathotype 3H was 3:1, which was also the most likely fit for pathotypes 3D and 5X. This indicated that resistance to all three pathotypes in the DH population was controlled by two genes in duplicate action. Two QTLs, *Rcr10*^{ECD01} and *Rcr9*^{ECD01}, for resistance to pathotypes 3A, 3D, and 3H were identified, which was consistent with the genetic analysis of phenotype ratios. However, only one QTL, *Rcr9*^{ECD01}, was identified for resistance to 5X, although the segregation ratio was close to 3:1. This inconsistency merits further investigation.

In general, strong resistance to clubroot pathotypes is controlled by single dominant genes such as *Rcr1*–*Rcr7* (Chu et al., 2014; Yu et al., 2016; Huang et al., 2017, 2019; Yu et al., 2017; Dakouri et al., 2018; Chang et al., 2019; Karim et al., 2020). Two genes in duplicate action (*Rcr8* on chromosome A02 and *Rcr9* on chromosome A08 from *B. rapa* line T19) that conferred resistance to pathotype 5X were reported previously (Yu et al., 2017). Similarly, a previous study indicated that neither *Crr1* nor *Crr2* on their own conferred resistance to Japanese strain “Wakayama-01” of *P. brassica*; resistance was only expressed when resistance alleles were present at both loci (Suwabe et al., 2003). In this study, the QTL for resistance to pathotype 3A derived from ECD01 may behave similarly to *Crr1* and *Crr2*.

The number of SNP sites per chromosome is usually correlated with chromosome size in mapping populations (Yu et al., 2016) but was not correlated in this study. This unusual result likely occurred because the BC₂ donor plant used for microspore culture carried a large fragment of chromosome A01 originating from ECD01 but smaller fragments of the other chromosomes from ECD01.

In this study, the target region for *Rcr10*^{ECD01} was defined as 9.8–13.5 Mb of *B. napus* chromosome A03 using QTL analysis. A similar interval (9–14 Mb) for *Rcr10*^{ECD01} was obtained using

the identification of the PPV with BSA. The region for *Rcr10*^{ECD01} in *B. napus* was homologous to the 11.0–16.0 Mb region of A03 in the *B. rapa* “Chiifu” version 3.0. This was a distinct genetic region from *Rcr1*, *Rcr2*, *Rcr4*, and *Rcr5* for resistance to pathotypes of *P. brassica* (Figure 4). The genes *Rcr1*, *Rcr2*, and *Rcr4*, which confer resistance to pathotypes of *P. brassica*, have previously been mapped into chromosome A03 of *B. rapa* “Chiifu” version 3.0 at ~25 Mb region (Chu et al., 2014; Yu et al., 2016; Huang et al., 2017), while *Rcr5* was also mapped at ~24 Mb region in that chromosome (Huang et al., 2019; Figure 4). *Rcr1*, *Rcr2*, and *Rcr4* were subsequently co-localized with the cloned CR genes *CRA/CRb*^{kato} (Ueno et al., 2012; Hatakeyama et al., 2017), while *Rcr5* was located in a region close to *CRA/CRb*^{kato}. In addition, resistance genes *Rcr1*, *Rcr2*, *Rcr4*, and *Rcr5* were identified for resistance to pathotype 3H, not for 3A, 3D, or 5X. Several CR genes or QTLs, such as *PbBa3.2* (Chen et al., 2013), *CRd* (Pang et al., 2018), *Crr3* (Hirai et al., 2004), and *CRk* (Sakamoto et al., 2008) for resistance to clubroot strains collected from Japan and China, have been mapped into the regions different from *CRA/CRb*^{kato}. Similarly, *BraA.CR.c* for resistance was mapped into chromosome A03 in turnip cvs. ECD01, ECD02, and ECD04 (Hirani et al., 2018). The relationship of *Rcr10*^{ECD01} to these previously identified genes needs to be determined. Also, *CRb* was identified in a Chinese cabbage cv. “CR Shinki” was originally derived from ECD01 for resistance to *P. brassica* strains collected in South Korea (Piao et al., 2004). It was located in a genetic region close to *CRA/CRb*^{kato}. However, no QTL in the *CRb* region was identified in this study.

A QTL, identified and designated as *Rcr9*^{ECD01} (because it was mapped into the genetic region of *Rcr9* and was originally derived from *B. rapa* cv. ECD01), conferred resistance to all four pathotypes (3A, 3D, 3H, and 5X) assessed in this study. *Rcr9*^{ECD01} was located on the 2.2 Mb length from 9.3 to 11.5 Mb

of *B. napus* chromosome A08 using QTL analysis. The *Rcr9*^{ECD01} interval was confirmed through the identification of PPV with BSA, located in the physical interval 9–12 Mb. The region of *Rcr9*^{ECD01} in *B. napus* corresponded to 12.0–14.5 Mb of A08 in *B. rapa* “Chiifu” version 3.0 (Figure 4).

Previously, our laboratory had identified *Rcr9* for resistance to pathotype 5X in *B. rapa* breeding line T19, which originated from German turnip cv. “Pluto” (Yu et al., 2017). The proposed position of *Rcr9* spanned a large interval (6.48 Mb) of chromosome A08, including the genome region of *Rcr3* and *Rcr9*^{ECD01}. However, several breeding lines that carried *Rcr9* were resistant to 5X, but not to 3A, 3D (Yu, unpublished), and 3H (Yu et al., 2017). This difference in phenotype indicated that *Rcr9* differed from *Rcr9*^{ECD01}. Another resistance gene, designated as *Rcr9*^{wa}, has also been identified from a turnip differential line in the ECD. It originated from cv. “Waaslander” (ECD04), provided resistance to pathotype 5X, and was mapped into the same region as *Rcr9* (Karim et al., 2020). *Rcr9*^{wa} was mapped based on flanking markers into 12.3–12.6 Mb of chromosome A08 (smaller interval than *Rcr9*). In addition, another resistance gene originated from cv. “Waaslander” and conferred resistance to pathotype 3H, designated as *Rcr3*, has been mapped into chromosome A08, flanked by SNP markers in position 11.3–11.6 Mb in the *B. rapa* “Chiifu” reference genome version 3.0 (Karim et al., 2020). The position of *Rcr3* was separated from *Rcr9*^{ECD01} (Figure 4). Also, gene *BraA.CR.b* for resistance to pathotype 3H was previously identified from the turnip differentials ECD01, ECD02, ECD03, and ECD04 and mapped into chromosome A08 (Hirani et al., 2018), but no information on the genome region corresponding to the *B. rapa* “Chiifu” reference genome version 3.0 was provided. Several genes for resistance to collections of *P. brassicae* from Japan and China, including *Crr1* (Suwabe et al., 2003), *CRs* (Laila et al., 2019), *PbBa8.1* (Chen et al., 2013), and *qBrCR38-2* (Zhu et al., 2019), have also been mapped into chromosome A08. The cloned CR gene *Crr1* was highly homologous to *Bra020861* in the *B. rapa* reference genome version 1.5 and to *BraA08g014480* in the *B. rapa* reference genome version 3.0, which is located in the *Rcr9*^{ECD01} genomic region. However, breeding lines carrying *Crr1* gene did not show resistance to the strains of *P. brassicae* used in this study (Yu, unpublished). Therefore, *Rcr9*^{ECD01} is unlikely the same as *Crr1*. The relationship of *Rcr9*^{ECD01} with *CRs* (Laila et al., 2019), *PbBa8.1* (Chen et al., 2013), and *qBrCR38-2* needs to be determined.

CRc was identified in the Chinese cabbage cv. “CR Kanko” derived from “Debra,” which was located into chromosome A02 (Sakamoto et al., 2008). However, this gene was not found in the DH population used for this study.

Analysis of QTLs has been used for the identification of several major genes for resistance to clubroot (Yu et al., 2017). A QTL that can be consistently detected with a PVE of > 10% of trait value can be designated as the main effect QTL or major QTL (Wang et al., 2019). In this study, QTLs *Rcr10*^{ECD01} and *Rcr9*^{ECD01} were identified with 16.2 to 43.3% PVE. *Rcr10*^{ECD01} was identified based on the response to inoculation with pathotypes 3A, 3D, and 3H. *Rcr9*^{ECD01} was identified based on the response to inoculation with all of the pathotypes used in this study. Therefore, both *Rcr10*^{ECD01} and *Rcr9*^{ECD01} appear

to be major QTLs. The presence of the two major QTLs was also confirmed through BSA, which is consistent with the result obtained from the QTL analysis.

Clubroot severity in the DH lines in response to inoculation with the individual pathotypes was highly correlated, which indicated that resistance to these pathotypes was likely controlled by the same gene or tightly linked genes. However, the identification of QTLs in this study was based on relatively rough gene mapping, so it could not be determined if resistance to the pathotypes was controlled by a single gene or tightly linked genes. More detailed studies are in progress.

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.

AUTHOR CONTRIBUTIONS

FY conceived the study, designed the experiments, performed the data analysis, and drafted the manuscript. YZ and JW developed the population, performed the data analysis, and collected phenotypic data. QC and MK performed the data analysis. BG and GP provided important materials. All authors reviewed the manuscript and approved the final draft.

FUNDING

This study was funded by the Saskatchewan Agriculture Development Fund, SaskCanola, and Western Grains Research Foundation.

ACKNOWLEDGMENTS

The authors thank S. E. Strelkov for providing the strains of *P. brassicae*, Nutrien for the *B. rapa* donor ECD01, G. Séguin-Swartz for the recipient *B. napus* line DH16516, and M. Kehler and L. Liu for technical assistance with plant phenotyping.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.785989/full#supplementary-material>

Supplementary Figure 1 | The total numbers of sequences obtained from genotype by sequencing and the number of sequence aligned into the A-genome of *B. napus* “Darmor-bzh” version 4.1 for each line in the DH population.

Supplementary Figure 2 | The linkage map of A-genome chromosomes of *B. napus* developed from 260 SNP sites. Note that the A06 chromosome did not contain any SNPs and so is not represented.

REFERENCES

- Botero-Ramírez, A., Laperche, A., Guichard, S., Jubault, M., Grivot, A., Strelkov, S. E., et al. (2020). Clubroot symptoms and resting spore production in a doubled haploid population of Oilseed Rape (*Brassica napus*) are controlled by four main QTLs. *Front. Plant Sci.* 11:1998. doi: 10.3389/fpls.2020.604527
- Buczacki, S., Toxopeus, H., Mattusch, P., Johnston, T., Dixon, G., and Hobolth, L. (1975). Study of physiologic specialization in *Plasmodiophora brassicae*: proposals for attempted rationalization through an international approach. *Trans. Br. Mycol. Soc.* 65, 295–303.
- Chang, A., Lamara, M., Wei, Y., Hu, H., Parkin, I. A., Gossen, B. D., et al. (2019). Clubroot resistance gene Rcr6 in *Brassica nigra* resides in a genomic region homologous to chromosome A08 in *B. rapa*. *BMC Plant Biol.* 19:224. doi: 10.1186/s12870-019-1844-5
- Chen, J., Jing, J., Zhan, Z., Zhang, T., Zhang, C., and Piao, Z. (2013). Identification of novel QTLs for isolate-specific partial resistance to *Plasmodiophora brassicae* in *Brassica rapa*. *PLoS One* 8:e85307. doi: 10.1371/journal.pone.0085307
- Chu, M., Song, T., Falk, K. C., Zhang, X., Liu, X., Chang, A., et al. (2014). Fine mapping of Rcr1 and analyses of its effect on transcriptome patterns during infection by *Plasmodiophora brassicae*. *BMC Genomics* 15:1166. doi: 10.1186/1471-2164-15-1166
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- Dakouri, A., Lamara, M., Karim, M. M., Wang, J., Chen, Q., Gossen, B. D., et al. (2021). Identification of resistance loci against new pathotypes of *Plasmodiophora brassicae* in *Brassica napus* based on genome-wide association mapping. *Sci. Rep.* 11:6599. doi: 10.1038/s41598-021-85836-9
- Dakouri, A., Zhang, X., Peng, G., Falk, K. C., Gossen, B. D., Strelkov, S. E., et al. (2018). Analysis of genome-wide variants through bulked segregant RNA sequencing reveals a major gene for resistance to *Plasmodiophora brassicae* in *Brassica oleracea*. *Sci. Rep.* 8:17657. doi: 10.1038/s41598-018-36187-5
- Diederichsen, E., Frauen, M., Linders, E. G., Hatakeyama, K., and Hirai, M. (2009). Status and perspectives of clubroot resistance breeding in crucifer crops. *J. Plant Growth Regul.* 28, 265–281. doi: 10.1007/s00344-009-9100-0
- Fredua-Agyeman, R., and Rahman, H. (2016). Mapping of the clubroot disease resistance in spring *Brassica napus* canola introgressed from European winter canola cv. 'Mendel'. *Euphytica* 211, 201–213.
- Gowers, S. (1982). The transfer of characters from *Brassica campestris* L. to *Brassica napus* L.: production of clubroot-resistant oil-seed rape (*B. napus* ssp. *oleifera*). *Euphytica* 31, 971–976. doi: 10.1007/bf00039237
- Hasan, M. J., and Rahman, H. (2016). Genetics and molecular mapping of resistance to *Plasmodiophora brassicae* pathotypes 2, 3, 5, 6, and 8 in rutabaga (*Brassica napus* var. *napobrassica*). *Genome* 59, 805–815. doi: 10.1139/gen-2016-0034
- Hatakeyama, K., Niwa, T., Kato, T., Ohara, T., Kakizaki, T., and Matsumoto, S. (2017). The tandem repeated organization of NB-LRR genes in the clubroot-resistant CRb locus in *Brassica rapa* L. *Mol. Genet. Genom.* 292, 397–405. doi: 10.1007/s00438-016-1281-1
- Hatakeyama, K., Suwabe, K., Tomita, R. N., Kato, T., Nunome, T., Fukuoka, H., et al. (2013). Identification and characterization of Crr1a, a gene for resistance to clubroot disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *PLoS One* 8:e54745. doi: 10.1371/journal.pone.0054745
- Hirai, M. (2006). Genetic analysis of clubroot resistance in *Brassica* crops. *Breed. Sci.* 56, 223–229. doi: 10.1270/jsbbs.56.223
- Hirai, M., Harada, N., Kubo, N., Tsukada, M., Suwabe, K., and Matsumoto, S. (2004). A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. *Theor. Appl. Genet.* 108, 639–643. doi: 10.1007/s00122-003-1475-x
- Hirani, A. H., Gao, F., Liu, J., Fu, G., Wu, C., Mcvetty, P. B., et al. (2018). Combinations of independent dominant loci conferring clubroot resistance in all four turnip accessions (*Brassica rapa*) from the European clubroot differential set. *Front. Plant Sci.* 9:1628. doi: 10.3389/fpls.2018.01628
- Hollman, K., Hwang, S., Manoli, V., and Strelkov, S. (2021). Pathotypes of *Plasmodiophora brassicae* collected from clubroot resistant canola (*Brassica napus* L.) cultivars in western Canada in 2017–2018. *Can. J. Plant Pathol.* 43, 622–630.
- Horiuchi, S., and Hori, M. (1980). A simple greenhouse technique for obtaining high levels of clubroot [disease] incidence [of Brassica vegetables]. *Bull. Chugoku Natl. Agric. Exp. Stn. Ser. E. Environ. Div.* 17, 33–55.
- Huang, Z., Peng, G., Gossen, B., and Yu, F. (2019). Fine mapping of a clubroot resistance gene from turnip using SNP markers identified from bulked segregant RNA-Seq. *Mol. Breed.* 39, 1–10. doi: 10.3389/fpls.2017.01448
- Huang, Z., Peng, G., Liu, X., Deora, A., Falk, K. C., Gossen, B. D., et al. (2017). Fine mapping of a clubroot resistance gene in Chinese cabbage using SNP markers identified from bulked segregant RNA sequencing. *Front. Plant Sci.* 8:1448.
- Iversen, G. R., and Gergen, M. (1997). *Statistics: The Conceptual Approach*. New York, NY: Springer Science & Business Media.
- Johnston, T. (1974). Transfer of disease resistance from *Brassica campestris* L. to rape (*B. napus* L.). *Euphytica* 23, 681–683.
- Karim, M., Dakouri, A., Zhang, Y., Chen, Q., Peng, G., Strelkov, S. E., et al. (2020). Two clubroot-resistance genes, Rcr3 and Rcr9wa, mapped in *Brassica rapa* using bulk segregant RNA sequencing. *Int. J. Mol. Sci.* 21:5033. doi: 10.3390/ijms21145033
- Laila, R., Park, J.-I., Robin, A. H. K., Natarajan, S., Vijayakumar, H., Shirasawa, K., et al. (2019). Mapping of a novel clubroot resistance QTL using ddRAD-seq in Chinese cabbage (*Brassica rapa* L.). *BMC Plant Biol.* 19:13. doi: 10.1186/s12870-018-1615-8
- Lammerink, J. (1970). Inter-specific transfer of clubroot resistance from *Brassica campestris* L. to *B. napus* L. *N. Z. J. Agric. Res.* 13, 105–110.
- Lee, J., Izzah, N. K., Choi, B.-S., Joh, H. J., Lee, S.-C., Perumal, S., et al. (2016). Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Res.* 23, 29–41. doi: 10.1093/dnares/dsv034
- Manzanares-Dauleux, M., Delourme, R., Baron, F., and Thomas, G. (2000). Mapping of one major gene and of QTLs involved in resistance to clubroot in *Brassica napus*. *Theor. Appl. Genet.* 101, 885–891. doi: 10.1007/s001220051557
- Matsumoto, E., Yasui, C., Ohi, M., and Tsukada, M. (1998). Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Euphytica* 104, 79–86.
- 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
- Michelmores, R. W., Paran, I., and Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *PNAS* 88, 9828–9832. doi: 10.1073/pnas.88.21.9828
- Moringa, T. (1934). Interspecific hybridization in *Brassica* VI. The cytology of F 1 hybrids of *B. juncea* and *B. nigra*. *Cytologia* 6, 62–67. doi: 10.1508/cytologia.6.62
- Nagaharu, U., and Nagaharu, N. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Bot.* 7, 389–452.
- Nikolaev, S. I., Berney, C., Fahrni, J.-F., Bolivar, I., Polet, S., Mylnikov, A. P., et al. (2004). The twilight of *Heliozoa* and rise of *Rhizaria*, an emerging supergroup of amoeboid eukaryotes. *Proc. Natl. Acad. Sci.* 101, 8066–8071. doi: 10.1073/pnas.0308602101
- Pageau, D., Lajeunesse, J., and Lafond, J. (2006). Impact de l'hernie des crucifères [*Plasmodiophora brassicae*] sur la productivité et la qualité du canola. *Can. J. Plant Pathol.* 28, 137–143. doi: 10.1080/07060660609507280
- Pang, W., Fu, P., Li, X., Zhan, Z., Yu, S., and Piao, Z. (2018). Identification and mapping of the clubroot resistance gene CRd in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Front. Plant Sci.* 9:653. doi: 10.3389/fpls.2018.00653
- Pang, W., Liang, S., Li, X., Li, P., Yu, S., Lim, Y. P., et al. (2014). Genetic detection of clubroot resistance loci in a new population of *Brassica rapa*. *Hortic. Environ. Biotechnol.* 55, 540–547.
- Peng, L., Zhou, L., Li, Q., Wei, D., Ren, X., Song, H., et al. (2018). Identification of quantitative trait loci for clubroot resistance in *Brassica oleracea* with the use of *Brassica* SNP microarray. *Front. Plant Sci.* 9:822. doi: 10.3389/fpls.2018.00822
- Piao, Z., Deng, Y., Choi, S., Park, Y., and Lim, Y. (2004). SCAR and CAPS mapping of CRb, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Theor. Appl. Genet.* 108, 1458–1465. doi: 10.1007/s00122-003-1577-5

- Piao, Z., Ramchiary, N., and Lim, Y. P. (2009). Genetics of clubroot resistance in *Brassica* species. *J. Plant Growth Regul.* 28, 252–264.
- Sakamoto, K., Saito, A., Hayashida, N., Taguchi, G., and Matsumoto, E. (2008). Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Theor. Appl. Genet.* 117, 759–767. doi: 10.1007/s00122-008-0817-0
- Strelkov, S. E., Hwang, S.-F., Manolii, V. P., Cao, T., Fredua-Agyeman, R., Harding, M. W., et al. (2018). Virulence and pathotype classification of *Plasmodiophora brassicae* populations collected from clubroot resistant canola (*Brassica napus*) in Canada. *Can. J. Plant Pathol.* 40, 284–298.
- Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Fujimura, M., Nunome, T., et al. (2003). Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *Theor. Appl. Genet.* 107, 997–1002. doi: 10.1007/s00122-003-1309-x
- Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fujimura, M., et al. (2006). Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* 173, 309–319. doi: 10.1534/genetics.104.038968
- Ueno, H., Matsumoto, E., Aruga, D., Kitagawa, S., Matsumura, H., and Hayashida, N. (2012). Molecular characterization of the CRa gene conferring clubroot resistance in *Brassica rapa*. *Plant Mol. Biol.* 80, 621–629. doi: 10.1007/s11103-012-9971-5
- Van Ooijen, J. (2011). Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genetics Res.* 93, 343–349. doi: 10.1017/S0016672311000279
- Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77
- Voorrips, R. E. (1995). *Plasmodiophora brassicae*: aspects of pathogenesis and resistance in *Brassica oleracea*. *Euphytica* 83, 139–146.
- Wang, L., Yang, X., Cui, S., Mu, G., Sun, X., Liu, L., et al. (2019). QTL mapping and QTL × environment interaction analysis of multi-seed pod in cultivated peanut (*Arachis hypogaea* L.). *Crop J.* 7, 249–260. doi: 10.1016/j.cj.2018.11.007
- Werner, S., Diederichsen, E., Frauen, M., Schondelmaier, J., and Jung, C. (2008). Genetic mapping of clubroot resistance genes in oilseed rape. *Theor. Appl. Genet.* 116, 363–372. doi: 10.1007/s00122-007-0674-2
- Yu, F., Zhang, X., Huang, Z., Chu, M., Song, T., Falk, K. C., et al. (2016). Identification of genome-wide variants and discovery of variants associated with *Brassica rapa* clubroot resistance gene Rcr1 through bulked segregant RNA sequencing. *PLoS One* 11:e0153218. doi: 10.1371/journal.pone.0153218
- Yu, F., Zhang, X., Peng, G., Falk, K. C., Strelkov, S. E., and Gossen, B. D. (2017). Genotyping-by-sequencing reveals three QTL for clubroot resistance to six pathotypes of *Plasmodiophora brassicae* in *Brassica rapa*. *Sci. Rep.* 7:4516. doi: 10.1038/s41598-017-04903-2
- Zhang, L., Cai, X., Wu, J., Liu, M., Grob, S., Cheng, F., et al. (2018). Improved *Brassica rapa* reference genome by single-molecule sequencing and chromosome conformation capture technologies. *Hortic. Res.* 5, 1–11.
- Zhu, H., Zhai, W., Li, X., and Zhu, Y. (2019). Two QTLs controlling clubroot resistance identified from bulked segregant sequencing in Pakchoi (*Brassica campestris* ssp. *chinensis* Makino). *Sci. Rep.* 9:9228. doi: 10.1038/s41598-019-44724-z

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 Yu, Zhang, Wang, Chen, Karim, Gossen and Peng. 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 and Genetic Characterization of Peanut Advanced Backcross Lines That Incorporate Root-Knot Nematode Resistance From *Arachis stenosperma*

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

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

Phatu William Mashela,
University of Limpopo, South Africa

*Correspondence:

Soraya C. M. Leal-Bertioli
sbertioli@uga.edu

† Present address:

Carolina Ballén-Taborda,
Department of Plant
and Environmental Sciences,
Clemson University, Florence, SC,
United States

Scott A. Jackson,
Bayer Crop Science, Chesterfield,
MO, United States

Specialty section:

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

Received: 29 September 2021

Accepted: 01 December 2021

Published: 17 January 2022

Citation:

Ballén-Taborda C, Chu Y,
Ozias-Akins P, Holbrook CC,
Timper P, Jackson SA, Bertioli DJ and
Leal-Bertioli SCM (2022)
Development and Genetic
Characterization of Peanut Advanced
Backcross Lines That Incorporate
Root-Knot Nematode Resistance
From *Arachis stenosperma*.
Front. Plant Sci. 12:785358.
doi: 10.3389/fpls.2021.785358

Carolina Ballén-Taborda^{1†}, Ye Chu², Peggy Ozias-Akins^{1,2}, C. Corley Holbrook³,
Patricia Timper³, Scott A. Jackson^{1†}, David J. Bertioli^{1,4} and Soraya C. M. Leal-Bertioli^{1,5*}

¹ Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, GA, United States, ² Department of Horticulture, University of Georgia, Tifton, GA, United States, ³ U.S. Department of Agriculture-Agricultural Research Service (USDA-ARS), Tifton, GA, United States, ⁴ Department of Crop and Soil Science, University of Georgia, Athens, GA, United States, ⁵ Department of Plant Pathology, University of Georgia, Athens, GA, United States,

Crop wild species are increasingly important for crop improvement. Peanut (*Arachis hypogaea* L.) wild relatives comprise a diverse genetic pool that is being used to broaden its narrow genetic base. Peanut is an allotetraploid species extremely susceptible to peanut root-knot nematode (PRKN) *Meloidogyne arenaria*. Current resistant cultivars rely on a single introgression for PRKN resistance incorporated from the wild relative *Arachis cardenasii*, which could be overcome as a result of the emergence of virulent nematode populations. Therefore, new sources of resistance may be needed. Near-immunity has been found in the peanut wild relative *Arachis stenosperma*. The two loci controlling the resistance, present on chromosomes A02 and A09, have been validated in tetraploid lines and have been shown to reduce nematode reproduction by up to 98%. To incorporate these new resistance QTL into cultivated peanut, we used a marker-assisted backcrossing approach, using PRKN A. *stenosperma*-derived resistant lines as donor parents. Four cycles of backcrossing were completed, and SNP assays linked to the QTL were used for foreground selection. In each backcross generation seed weight, length, and width were measured, and based on a statistical analysis we observed that only one generation of backcrossing was required to recover the elite peanut's seed size. A population of 271 BC₃F₁ lines was genome-wide genotyped to characterize the introgressions across the genome. Phenotypic information for leaf spot incidence and domestication traits (seed size, fertility, plant architecture, and flower color) were recorded. Correlations between the wild introgressions in different chromosomes and the phenotypic data allowed us to identify candidate regions controlling these domestication traits. Finally, PRKN resistance was validated in BC₃F₃ lines. We observed that the QTL in A02 and/or large introgression in A09 are needed for resistance. This present work represents an important step toward the development of new high-yielding and nematode-resistant peanut cultivars.

Keywords: wild crop relatives, *Arachis*, peanut, root-knot nematode, *Meloidogyne arenaria*, marker-assisted backcrossing, domestication

INTRODUCTION

Arachis hypogaea L., with a common name of peanut or groundnut, is an important oil, food, and fodder crop cultivated worldwide with an annual production of 66.3 million tons and grown on 34.1 Mha (FAOSTAT, 2021). Peanut is an allotetraploid species (AABB, $2n = 4x = 40$), with a recent and unique polyploid origin, which occurred 5 to 10 thousand years ago (Bertioli et al., 2019, 2020). This narrow genetic base and limited gene flow with its genetically diverse diploid wild relatives resulted in a lack of strong resistance alleles for pests and diseases in the primary gene pool. One important pest is the peanut root-knot nematode (PRKN) (*Meloidogyne arenaria*) (Holbrook and Stalker, 2003). It causes yield losses greater than 50% in infested fields, and at times, 100% losses in heavily infested areas of fields have been reported (Dickson and De Waele, 2005; Timper et al., 2018). In the United States, *M. arenaria* is the most damaging nematode for peanut (Timper et al., 2018). Chemical control is one option, but is costly, hazardous to human health, and can damage the environment (Oka, 2020). Crop rotation is also effective, but with susceptible cultivars, the constraints on the frequency at which peanut can be grown reduce agronomic and financial sustainability. The use of high-yielding and nematode-resistant cultivars in combination with rotation is the most efficient and effective way to control nematode populations and maintain yield while reducing the use of nematicides.

Strong resistance to many pests and diseases is limited in the *A. hypogaea* primary gene pool (Stalker, 2017), which imposes constraints for crop improvement using cultivated germplasm (Nelson et al., 1989; Noe et al., 1992). Yet, the wild relatives comprise a diverse genetic pool that has the potential to broaden a peanut's genetic base and to improve its performance under pest/disease pressure (Stalker, 2017). Previously, successful transfer of root-knot nematode resistance into cultivated peanut was accomplished through backcrossing schemes involving a synthetic allotetraploid (Simpson and Starr, 2001). This resistance is derived from introgression of a large segment on chromosome A09 from the wild species *Arachis cardenasii* (Nagy et al., 2010; Chu et al., 2016), and is present in several commercial cultivars (Georgia-14N, TifNV-High O/L, Tifguard, NemaTAM, and Webb) (Simpson et al., 2003, 2013; Holbrook et al., 2008, 2017; Branch and Brenneman, 2015). Although this resistance has been durable thus far, occasional resistance breakdown has been reported (Holbrook CC, personal communication). Therefore, it is important to incorporate new sources of resistance to reduce the risk of selection of a virulent population of *M. arenaria* and to guarantee continued protection of the peanut crop from losses due to PRKN.

The peanut wild relative *Arachis stenosperma* PI666100/V10309 has been described as highly resistant to peanut root-knot nematode (Proite et al., 2008). Previously, three quantitative trait loci (QTL) (on chromosomes A02, A04, and A09) were identified in the diploid genome of *Arachis stenosperma* (Leal-Bertioli et al., 2016). Later, segments of both chromosomes A02 and A09, that provide near immunity, were mapped using a segregating population derived from a cross between *A. hypogaea* and the synthetic allotetraploid

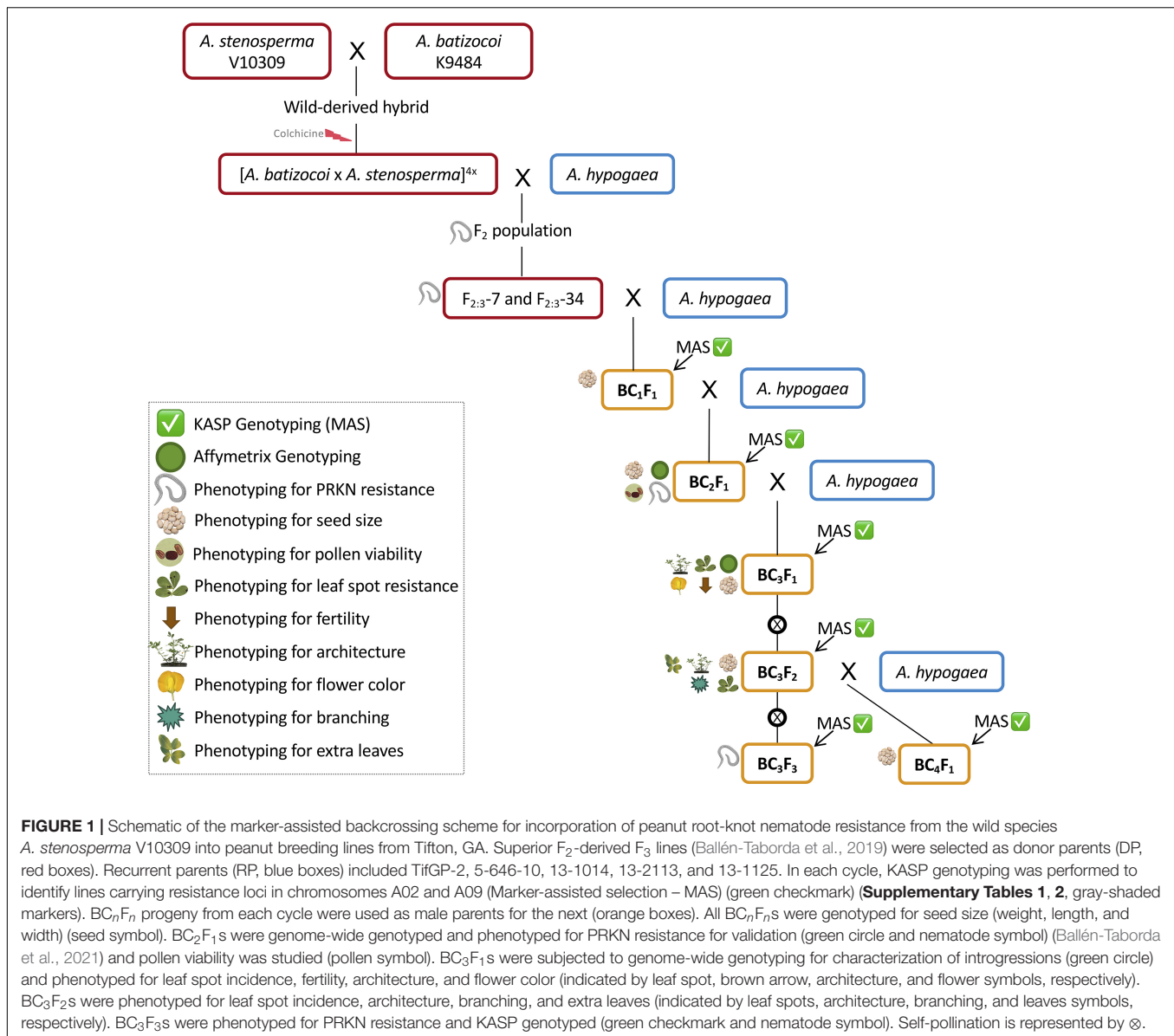
BatSten1 (Bertioli et al., 2021a), and validated in a tetraploid background (Ballén-Taborda et al., 2019, 2021). The main objective of this study was to incorporate PRKN resistance QTL from *A. stenosperma* into elite peanut. To accomplish this goal, a marker-assisted backcross breeding approach was employed and BatSten1 was used as the donor parent. Four cycles of backcrossing were completed with genetic foreground and background selection and phenotypic characterization were performed in each generation. Correlations between the wild introgressions across the genome and the phenotypic data allowed us to identify candidate regions controlling traits measured in the BC₃F₁ population.

This work is key to developing new high-yielding peanut cultivars with a new and strong resistance against the peanut root-knot nematode. Additionally, single nucleotide polymorphism (SNP) markers tightly linked to the QTL are described to facilitate the introgression of *A. stenosperma* resistance into different elite recipient lines. In the near future, we expect to release advanced introgression lines that incorporate strong PRKN resistance with attached molecular information, that can be used directly in breeding programs in areas where PRKN is a constraint for peanut cultivation.

MATERIALS AND METHODS

Plant Materials

The synthetic allotetraploid BatSten1 PI 695418 {[*Arachis batizocoi* PI298639/K9484 x *A. stenosperma* PI666100/V10309]^(2n = 4x = 40)} (Bertioli et al., 2021a) was used to introgress the nematode resistance QTL from *A. stenosperma* into tetraploid peanut. An F₂ population was created by selfing an F₁ derived from a cross between *A. hypogaea* cv. Runner IAC-886 (herein called Runner-886) and BatSten1, and then used for QTL mapping (Ballén-Taborda et al., 2019). From this population, four superior F₂ lines (F₂-7, F₂-13, F₂-34, and F₂-73) were selected based on (1) better vigor, as visually more leaf biomass; (2) good agronomic traits; (3) late leaf spot (LLS) and PRKN resistance; (4) harboring QTL in A02 and A09 for nematode resistance per molecular genotyping. Six F₂-derived F₃ (F₂:3) homozygous progeny from F₂:3-7 and F₂:3-34, with validated resistance to PRKN (Ballén-Taborda et al., 2021), were used as initial donor parents in a backcrossing scheme (Figure 1, red boxes). For incorporation of resistance from *A. stenosperma* into peanut, we used three susceptible recurrent parents, which are TifGP-2, 5-646-10, and 13-1014. For pyramiding resistances from both wild species *A. stenosperma* and *A. cardenasii*, we used two resistant lines, which are 13-2113 and 13-1125. (1) TifGP-2 is a breeding line with good yield and grade, and normal oleic content (Holbrook et al., 2012); (2) 5-646-10 is derived from the cross Florida-07 x Tifguard, with good yield and grade, and high oleic/linoleic fatty acid ratio (Holbrook, CC, unpublished data); (3) 13-1014 is derived from [C1805-617-1 (Florida-07 x Tifguard) x GA-06G], with high oleic/linoleic fatty acid ratio (Holbrook, CC, unpublished data); (4) 13-2113 is derived from [C1805-2-9-16 (Florida-07 x Tifguard) x TifGP-2], a high oleic/linoleic fatty acid ratio, that was included in the



second cycle only (Holbrook, CC, unpublished data); (5) 13-1125 breeding line was included in the fourth cycle only (Holbrook, CC, unpublished data) (Figure 1, blue boxes).

Marker-Assisted Breeding

A marker-assisted backcrossing (MABC) approach was used to incorporate PRKN resistance from *A. stenosperma* into cultivated peanut. Four cycles of backcrossing were performed in two different locations: Athens, GA and Tifton, GA under greenhouse conditions. In the first, second, and third cycles, 16 SNP markers linked to the QTL in A02 and A09 (Leal-Bertioli et al., 2016) were used for foreground selection (Supplementary Table 1, gray-shaded markers). For the fourth cycle, 10 new markers were used to finely target these chromosome segments based on high-throughput genotyping of backcross (BC) lines (Ballén-Taborda et al., 2021) (Supplementary Table 2,

gray-shaded markers). 60 additional SNPs markers were also developed and are available here for genotypic selection in breeding programs (Supplementary Tables 1, 2, not shaded markers). Progeny from each cycle that harbored segments associated with PRKN resistance were used as male parents for the next backcross cycle (Figure 1, orange boxes). BC_4F_1 s were germinated for seed increase.

Kompetitive Allele-Specific PCR assays (KASP, LGC Biosearch Technologies, Hoddesdon, United Kingdom)¹ assays were used for the selection of PRKN resistance alleles. For KASP reactions, genomic DNA was isolated from a small section of the peanut cotyledon opposite to the embryo (seed chip, ~50-100 mg). DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN,

¹<https://biosearch-cdn.azureedge.net/assetsv6/KASP-genotyping-chemistry-User-guide.pdf>

Hilden, DE) according to the manufacturer's instructions. Each KASP marker consisted of three primers per SNP position (two allele-specific and one common flanking primer). Primers were designed using the web-based program BatchPrimer3² (USDA-ARS, Albany, CA, United States) (You et al., 2008) with the "Allele-specific primers and allele flanking primers" option. Parameters used were 60–120 bp in fragment size, GC content of 30–80%, and T_m between 58 and 60°. KASP primer assay mix per SNP position consisted of 12 µl (100 µM) of each allele-specific primer, 30 µl (100 µM) of the flanking primer, and 46 µl of H₂O. Single KASP reactions (5 µl) consisted of 2.5 µl of KASP 2x Master Mix (Low Rox 5000 V4.0), 0.07 µl of KASP primer assay mix, 1.93 µl of water, and 0.5 µl of DNA (10 ng/µl). Two replicates per primer per sample were included in each reaction, as well as no-template controls (NTCs). A C100 touch Thermal Cycler (BIO-RAD, Hercules, CA, United States) was used with the following conditions: 94°C for 15 min; 8 cycles of 94°C for 20 s and touchdown starting at 61°C for 1 min (dropping 6° per cycle); 31 cycles of 94°C for 10 s and 55°C for 1 min; 9 cycles of 94°C for 20 s, and 57°C for 1 min; 4°C hold. Fluorescence was read with a LightCycler®480 Instrument II (Roche Life Science, Switzerland) and analyzed using the LightCycler®480 Software (v.1.5.1.62) (Roche Life Science, Switzerland). Finally, data were exported into Microsoft Excel for analysis.

Genome-Wide Genotyping of BC₃F₁s

Genomic DNA of 271 BC₃F₁ lines and controls (BatSten1, Runner-886, 5-646-10, 13-1014, TifGP-2, Tifguard, and Tifrunner) were extracted from lyophilized leaves using the DNeasy 96 Plant Kit (QIAGEN, Hilden, DE, United States) according to the manufacturer's instructions. DNAs were genotyped with the Axiom_Arachis2 SNP array (Clevenger et al., 2018; Korani et al., 2019). Genotypic data was extracted and processed using the Axiom™ Analysis Suite software (v.4.0.3.3) (ThermoFisher Scientific, Waltham, MA, United States). Output was analyzed using custom shell scripts (see below) and resulting data were visualized as a color map in Microsoft Excel. The physical positions of the A-genome markers were determined according to the position of their orthologs in the *A. duranensis* pseudomolecules and the K-genome markers based on the *A. ipaensis* pseudomolecules (Bertioli et al., 2016).

The strategy to identify polymorphic SNP markers included two main steps. First, a set of polymorphic SNP markers between parental genotypes (BatSten1 ≠ Runner-886) was extracted. Original genotyping calls were replaced by numbers ("1" for BatSten, "2" for Runner-886, and "3" for a different genotype). Second, SNPs present in the genetic map previously identified (*A. stenosperma*-specific and *A. batizocoi*-specific markers) (Ballén-Taborda et al., 2019) were retrieved (Supplementary Script 1). Genotypic data were used to perform the principal component analysis (PCA) using the "dist" function in R.

Peanut Root-Knot Nematode Resistance Validation Using BC₃F₃s

BC₃F₃ segregating lines from five BC₃F₁ families and controls were evaluated for PRKN resistance to further validate the QTL in A02 and/or A09 (Table 1). The experiment included 72 BC₃F₃ lines, which were selected based on the genotypic information of the BC₃F₁ generation, by focusing on high cultivar genome recovery (89.1–95.9%) and with superior field performance (data not shown) of the BC₃F₂ generation during summer 2020. Selected lines included two BC₃F₃ lineages with the A02-QTL, two with the A09-QTL, one with both A02-QTL and A09-QTL, and one with a large A09-QTL. The synthetic allotetraploid BatSten1 and the cultivar *A. hypogaea* TifNV-High O/L (Holbrook et al., 2017) were used as resistant controls and *A. hypogaea* 5-646-10 and 13-1014 as susceptible controls. To confirm the presence of the QTL, BC₃F₃s and controls were genotyped using KASP markers (as described in the "Marker-assisted breeding" section). Six KASP markers targeting the bottom of A02 (A02-83,464,195, A02-92,077,207, and A02-92,983,792) and A09 (A09-16,516,448, A09-112309,231 and A09-114,515,959) (Supplementary Table 2) were used. Additionally, Axiom_Arachis2 SNP array genotyping was performed for genome-wide characterization, and data were filtered as described above.

Peanut root-knot nematode (PRKN) populations were cultured and extracted from eggplant (*Solanum melongena*) to be used as inoculum for bioassays. Second stage juveniles (J₂s) were collected from infected roots in a mist chamber every 2–3 days over a week and stored at 10°C until inoculation. Peanut seeds were grown in nursery pots filled with steam-sterilized sandy soil in the greenhouse. Bioassay was performed under greenhouse conditions, in a randomized complete block design with 12 replicates per genotype. Furthermore, 40-day-old plants were inoculated with 6,000 J₂s by adding the inoculum in two 2-cm deep holes at the base of the plant. Two months later, plants were uprooted, rinsed to remove soil, assessed for galling, and weighed. Eggs were extracted from roots using 0.5% NaOCl and counted (Hussey and Barker, 1973; Holbrook et al., 2003). Two different traits were measured: (1) galling index (GI), where 0 = no galls, 1 = 1–2 galls, 2 = 3–10 galls, 3 = 11–30 galls, 4 = 31–100 galls and 5 = more than 100 galls (Taylor and Sasser, 1978) and (2) number of eggs. Galling index and number of eggs per root weight (GI/g and eggs/g) were used for resistance assessment. A highly resistant plant was defined as such when the reproduction of nematodes was less than 20% of the reproduction in a susceptible plant (Taylor and Sasser, 1978).

Phenotypic Characterization Seed Size

In each cycle of backcrossing, 11 BC₁F₁, 30 BC₂F₁, 253 BC₃F₁, 101 BC₃F₂ and 25 BC₄F₁ seeds were phenotyped for weight (g), length (mm, longest point) and width (mm, widest point) prior to planting. For controls (*A. stenosperma* V10309; *A. batizocoi* K9484; BatSten1; *A. hypogaea* genotypes, 5-646-10, 13-1014, TifGP-2 and Runner-886), 10 individual seeds were measured (Table 2).

²<https://wheat.pw.usda.gov/demos/BatchPrimer3/>

TABLE 1 | Average and standard deviation for GI/g and eggs/g measured in BC₃F₃ lines, resistant, and susceptible controls using a pot bioassay.

| Genotype (short) | Type | PRKN resistance segment* | GI/g (avg ± SD) | Eggs/g (avg ± SD) |
|---|---|--|------------------|-----------------------|
| BatSten1 | Resistant control | All segments | 0.00 ± 0.00 (d) | 0.00 ± 0.00 (b) |
| <i>A. hypogaea</i> TifNV-High O/L (TifNV-O/L) | Resistant control** | - | 0.03 ± 0.08 (cd) | 6.66 ± 17.43 (b) |
| <i>A. hypogaea</i> 5-646-10 (5-646-10) | Susceptible control | - | 0.31 ± 0.26 (ab) | 579.18 ± 855.26 (ac) |
| <i>A. hypogaea</i> 13-1014 (13-1014) | Susceptible control | - | 0.42 ± 0.11 (b) | 1268.64 ± 1046.49 (a) |
| 14_BC ₃ F ₃ :C2633-2_3(16) S14_B (BC ₃ F ₃ _14_B) | Backcross line – BC ₃ F ₃ | Bottom A02 (A02) | 0.00 ± 0.00 (cd) | 30.71 ± 75.23 (b) |
| 170_BC ₃ F ₃ :BRD_C0050_Seed9_S77_A (BC ₃ F ₃ _170_A) | Backcross line – BC ₃ F ₃ | Bottom A02 (A02) | 0.02 ± 0.04 (cd) | 11.89 ± 25.08 (b) |
| 133_BC ₃ F ₃ :RBS_sd20_S3_01 (BC ₃ F ₃ _133) | Backcross line – BC ₃ F ₃ | Bottom A09 (A09) | 0.35 ± 0.18 (ab) | 1411.04 ± 1076.60 (a) |
| 135_BC ₃ F ₃ :RBS_sd20_S3_03 (BC ₃ F ₃ _135) | Backcross line – BC ₃ F ₃ | Bottom A09 (A09) | 0.25 ± 0.17 (a) | 946.29 ± 888.79 (a) |
| 202_BC ₃ F ₃ :RBS_sd14_S1_B (BC ₃ F ₃ _202_B) | Backcross line – BC ₃ F ₃ | Bottom Small A02 (A02-) and Bottom A09 (A09) | 0.13 ± 0.19 (ac) | 34.92 ± 73.71 (bc) |
| 203_BC ₃ F ₃ :RBS_sd14_S2_A (BC ₃ F ₃ _203_A) | Backcross line – BC ₃ F ₃ | Large A09 (A09+) | 0.09 ± 0.13 (ac) | 9.87 ± 27.91 (b) |

avg ± SD, average ± standard deviation. GI/g, galling index per gram of root. Eggs/g, number of eggs per gram of root. Columns with the same letter do not differ significantly ($P < 0.05$). Full data in **Supplementary Table 8**. *Resistance was derived from *A. stenosperra*, based on genotyping by Axiom_Arachis2 SNP array.

** Resistance was derived from *A. cardenasii*.

TABLE 2 | Average and standard deviation for seed weight, length, and width measured in BC_nF_n lines, diploid wild species, induced allotetraploid BatSten1, and cultivated controls.

| Genotype | Type | Weight (g) min – max (Avg ± SD) | Length (mm) min – max (Avg ± SD) | Width (mm) min – max (Avg ± SD) | N |
|----------------------------------|-------------------------------|---------------------------------|----------------------------------|---------------------------------|-----|
| <i>A. stenosperra</i> V10309 | Diploid wild species | 0.16 – 0.19 (0.17 ± 0.01 b) | 11.10 – 13.39 (11.86 ± 0.64 d) | 4.93 – 5.57 (5.21 ± 0.21 c) | 10 |
| <i>A. batizocoi</i> K9484 | Diploid wild species | 0.14 – 0.31 (0.22 ± 0.05 c) | 11.26 – 14.70 (12.95 ± 1.21 e) | 5.01 – 6.37 (5.70 ± 0.43 d) | 10 |
| BatSten1 | Induced allotetraploid | 0.13 – 0.20 (0.17 ± 0.03 b) | 11.18 – 14.45 (12.52 ± 0.92 e) | 4.34 – 5.41 (5.05 ± 0.33 c) | 10 |
| <i>A. hypogaea</i> 5-646-10 | Cultivated (Recurrent parent) | 0.67 – 1.03 (0.81 ± 0.10 a) | 14.30 – 20.22 (17.21 ± 1.87 ab) | 9.71 – 11.20 (10.43 ± 0.51 ab) | 10 |
| <i>A. hypogaea</i> 13-1014 | Cultivated (Recurrent parent) | 0.58 – 0.97 (0.84 ± 0.11 a) | 15.30 – 20.19 (18.04 ± 1.62 a) | 8.50 – 11.54 (10.64 ± 0.89 a) | 10 |
| <i>A. hypogaea</i> TifGP-2 | Cultivated (Recurrent parent) | 0.59 – 0.91 (0.75 ± 0.12 a) | 14.23 – 17.66 (16.48 ± 1.22 bc) | 9.06 – 10.89 (9.81 ± 0.63 b) | 10 |
| <i>A. hypogaea</i> Runner-886 | Cultivated control | 0.51 – 0.94 (0.74 ± 0.17 a) | 12.61 – 17.77 (15.07 ± 1.70 c) | 8.40 – 11.26 (10.23 ± 1.01 ab) | 10 |
| BC ₁ F ₁ s | Backcross lines | 0.29 – 1.46 (0.68 ± 0.34) | – | – | 11 |
| BC ₂ F ₁ s | Backcross lines | 0.25 – 1.47 (0.72 ± 0.24) | 11.01 – 21.46 (15.66 ± 2.15) | 7.48 – 13.91 (9.87 ± 1.34) | 30 |
| BC ₃ F ₁ s | Backcross lines | 0.09 – 1.31 (0.66 ± 0.25) | 9.22 – 20.83 (15.60 ± 2.30) | 3.71 – 13.42 (9.51 ± 2.04) | 253 |
| BC ₃ F ₂ s | Backcross lines | 0.40 – 1.17 (0.72 ± 0.16) | 13.47 – 22.05 (16.94 ± 1.93) | 8.19 – 12.57 (10.11 ± 0.94) | 101 |
| BC ₄ F ₁ s | Backcross lines | 0.42 – 0.97 (0.69 ± 0.13) | 12.95 – 17.33 (15.83 ± 1.09) | 7.96 – 12.32 (10.23 ± 1.06) | 25 |

Minimum and maximum values for controls and BC_nF₁ seeds are presented.

Weight, length, and width for controls with the same letter do not differ significantly ($P < 0.05$). Min-max (avg ± SD), minimum – maximum (average ± standard deviation) values are presented. N, number of seeds. Full data is in **Supplementary Table 9**.

Pollen Viability (BC₂F₁s)

Pollen viability (PV) was evaluated for the BC₂F₁s and controls (**Table 3**). Flowers were collected early morning (between 8:00 and 10:00 am) and stained with acetocarmine (Heslop-Harrison, 1992). Stained pollen grains were observed and counted under a microscope (40X). Pollen viability from 10 individual flowers (reps) per genotype was assessed as the percentage of stained pollen grains (Gaaliche et al., 2013).

Leaf Spot Incidence, Fertility, Architecture, and Flower Color (BC₃F₁s)

While the BC₃F₁s were growing in the greenhouse, segregation for different traits was noticed, including foliar disease incidence, fertility (number of pegs), plant architecture, and flower color (**Figures 2A–C**). Single plant measurements were recorded. Leaf spot incidence was scored as a categorical variable as: “yes” (1) for *A. hypogaea* phenotype (susceptible) or “no” (0) for the resistant phenotype (**Figure 2A**). A total number of pegs was counted for

assessment of fertility. Plant architecture or growth habit was scored from 1 to 4, with 1 being erect and 4 for prostrate growth habit (**Figure 2B**) (Pittman, 1995). Lastly, flower color was scored visually as orange (1) for *A. hypogaea* phenotype versus yellow (0) for the wild phenotype (**Figure 2C**).

Leaf Spot Incidence, Architecture, Branching, and Extra Leaves (BC₃F₂s)

A group of 101 BC₃F₂ plants from 12 BC₃F₁ families with a high recurrent parent (*A. hypogaea*) genome recovery (88.3–97.3%) were selected to create a group of lines carrying different sizes of introgressions in A02 and A09 for future gene cloning experiments. KASP assays targeting ten SNPs at A02-QTL and A09-QTL (**Supplementary Table 2**, gray-shaded markers) were used to confirm the presence of the QTL. While the plants were growing, several traits were recorded: leaf spot incidence and plant architecture were scored as above, in addition, a score of 5 for dwarf phenotype was included (**Figures 2A,B**). Branching was scored from an abnormal number of branches (high) (1)

TABLE 3 | Average and standard deviation for pollen viability (%) and the number of pods quantified in BC₂F₁s lines and controls.

| Genotype (short) | Type | Pollen viability (%) | Number of pods |
|---|---|----------------------|----------------|
| <i>A. stenosperma</i> V10309 | Diploid wild species | 94.90 ± 1.23 (b) | 107 |
| <i>A. batizocoi</i> K9484 | Diploid wild species | 71.69 ± 11.45 (cd) | 260 |
| BatSten1 | Induced allotetraploid | 75.51 ± 11.01 (c) | 172 |
| <i>A. hypogaea</i> 5-646-10 | Cultivated (Recurrent parent) | 93.63 ± 1.17 (ab) | 120 |
| <i>A. hypogaea</i> 13-1014 | Cultivated (Recurrent parent) | 90.24 ± 2.78 (ef) | 95 |
| BC ₂ F ₁ :C2633-2_3(16) | Backcross line – BC ₂ F ₁ | 92.10 ± 3.46 (ae) | 60 |
| BC ₂ F ₁ _BRD_C0049_Seed1 (BC ₂ F ₁ _Seed1) | Backcross line – BC ₂ F ₁ | 86.62 ± 5.60 (fg) | 109 |
| BC ₂ F ₁ _BRD_C0049_Seed2 (BC ₂ F ₁ _Seed2) | Backcross line – BC ₂ F ₁ | 79.50 ± 4.81 (c) | 48 |
| BC ₂ F ₁ _BRD_C0049_Seed7 (BC ₂ F ₁ _Seed7) | Backcross line – BC ₂ F ₁ | 80.84 ± 7.05 (cg) | 5 |
| BC ₂ F ₁ _BRD_C0049_Seed8 (BC ₂ F ₁ _Seed8) | Backcross line – BC ₂ F ₁ | 65.08 ± 6.05 (d) | 47 |
| BC ₂ F ₁ _BRD_C0050_Seed9 (BC ₂ F ₁ _Seed9) | Backcross line – BC ₂ F ₁ | 75.27 ± 5.19 (c) | 103 |
| BC ₂ F ₁ _BRD_C0055_Seed15 (BC ₂ F ₁ _Seed15) | Backcross line – BC ₂ F ₁ | Not evaluated | 49 |
| BC ₂ F ₁ _BRD_C0055_Seed17 (BC ₂ F ₁ _Seed17) | Backcross line – BC ₂ F ₁ | 65.58 ± 9.14 (d) | 107 |
| BC ₂ F ₁ _BRD_C0057_Seed28 (BC ₂ F ₁ _Seed28) | Backcross line – BC ₂ F ₁ | 89.81 ± 2.29 (ef) | 260 |
| BC ₂ F ₁ _BRD_C0058_Seed33 (BC ₂ F ₁ _Seed33) | Backcross line – BC ₂ F ₁ | 77.90 ± 3.68 (c) | 172 |

Pollen viability (%) – Percentage of viable pollen grains. The pollen viability (%) column with the same letter does not differ significantly ($P < 0.05$). Number of pods – Total number of pods per plant. The correlation between Pollen viability (%) and the Number of pods was -0.007 ($P < 0.05$).

to intermediate (2) to normal *A. hypogaea* phenotype (3) (see **Figure 2D**). Finally, extra leaves in leaflets were observed and scored as 1 for *A. hypogaea* and 0 for the presence of at least one extra leaf (**Figure 2E**).

Association Analysis Using the BC₃F₁ Population

Association between genome-wide genotypic data of BC₃F₁ lines and phenotypic data of seed weight, length and width, leaf spot incidence, fertility, architecture, and flower color were used to identify candidate wild introgressions that could be controlling these traits. For seed size, leaf spot incidence, fertility, and architecture a Pearson correlation was performed. Flower color was analyzed using a mixed linear model (MLM) in Tassel (v.5) (Ithaca, NY, United States) (Bradbury et al., 2007). A Manhattan plot was created in R using the “*qqman*” package (Turner, 2014) and thresholds calculated using the “*CalcThreshold*” package with the Bonferroni method (Hamazaki and Iwata, 2020).

Statistical Analysis

Phenotypic data for seed weight, length and width, pollen viability, and nematode resistance bioassay were analyzed using the package R. A Shapiro-Wilk test was used to test for normal distribution. Non-parametric Kruskal-Wallis one-way analysis of variance (Kruskal and Wallis, 1952) was used to evaluate differences at a 5% level of significance ($P < 0.05$). For the seed weight, length, and width (transformed to Log10 when needed) the Welch *t*-test was used to perform pair-wise comparisons between wild accessions, cultivated genotypes, and backcross generations. Additionally, the non-parametric Skillings-Mack test (Chatfield and Mander, 2009) was used to evaluate significant differences for the RCBD nematode bioassay ($P < 0.05$). Further analysis included the Wilcoxon signed-rank test for pairwise comparisons using false

discovery rate (FDR) correction to group samples by significant similarity ($P < 0.05$).

RESULTS

Marker-Assisted Breeding

Four generations of MABC for introgression of PRKN resistance from *A. stenosperma* were completed in two locations, Athens, GA, and Tifton, GA under greenhouse conditions (**Figure 1**). KASP genotyping was performed using 16 (for first, second and third cycles) and 10 SNP markers (fourth cycle) (**Supplementary Tables 1, 2**, gray-shaded markers). For the first cycle of backcrosses, 38 cross combinations were used, with 19 F_{2:3} plants as donor male parents and two cultivated peanut female parents, namely, TifGP-2 and 5-646-10. In this cycle, 1,008 potential BC₁F₁ seeds were harvested, and based on KASP genotyping, 17 were selected for harboring the nematode resistance segments in A02 and A09 (**Supplementary Table 3**).

For the second cycle, 14 cross combinations were made with four cultivated peanut female parents, TifGP-2, 5-646-10, 13-1014 and 13-2113, and 11 BC₁F₁ male parents. Here, 61 potential BC₂F₁ seeds were obtained and genotyped with KASP markers. A total of 21 were selected as they harbored resistance segments in A02 and A09 PRKN resistance QTL (**Supplementary Table 4**).

For the third cycle of backcrossing, a total of 21 cross combinations were made using 5-646-10 and 13-1014 as female parents and 21 BC₂F₁ male parents selected the previous year, six of which were shown to be resistant to PRKN (Ballén-Taborda et al., 2021). This resulted in 397 potential BC₃F₁ seeds (**Supplementary Table 5**), that were genotypically characterized in two different ways. First, 81 BC₃F₁ seeds were randomly selected and evaluated for the presence of resistance segments using the KASP assays. From this group, 52 BC₃F₁s harbored the segments in A02 and A09 from *A. stenosperma*



FIGURE 2 | Phenotypic characterization of advanced backcross population. Leaf spot incidence for BC₃F₁s and BC₃F₂s ["yes" (1) or "no" (0)] (A); architecture for BC₃F₁s and BC₃F₂s (scores: 1 – erect, 2 and 3 – intermediate, 4 – prostrate, and 5 – dwarf phenotype observed only for BC₃F₂ generation) (Pittman, 1995) (B); flower color for BC₃F₁s (scores: Orange – 1 or yellow – 0) (C); variation in branching for BC₃F₂s (scores: 1 – high, 2 – intermediate and 3 – normal) (D); 1–3 extra leaves in leaflets for BC₃F₂s (score 0) (E). "*" indicates *A. hypogaea* phenotype in all the cases.

(Supplementary Table 5). Second, 271 BC₃F₁s were genome-wide genotyped (see next section).

For the fourth cycle, six cross combinations were used with 5-646-10, 13-1014, and 13-1125 as females and two BC₃F₂ male parents harboring A02 and A09 loci. Here, 27 potential BC₄F₁ seeds were obtained and genotyped with KASP markers. A total of 25 were confirmed to harbor PRKN resistance chromosome segments at the bottom of both A02 and A09 (Supplementary Table 6). Three of them could have combined both sources of resistance in A02 from *A. stenosperma* and A09 from *A. cardenasii* present in parent 13-1125.

Genome-Wide Genotyping of BC₃F₁s

To characterize the wild introgressions in the BC₃F₁ population, 271 lines and controls were genotyped with the Axiom_Arachis2 SNP array. A total of 930 informative polymorphic SNP markers, previously identified and assigned to A and B subgenomes (Ballén-Taborda et al., 2019), were recovered. Among these, 527 markers were located in A-subgenome (*A. stenosperma*-specific markers) and 403 to B/K-subgenome (*A. batizocoi*-specific markers). Of the 271 genotyped lines, 253 (93.4%) were true progeny from hybridization and 18 (6.6%) were products of self-pollination (Supplementary Table 7).

Examples of Axiom clustering plots of SNPs linked to QTL in A02 (Supplementary Figures 1A,B) and A09 (Supplementary Figures 1C–E) show the distribution of BC₃F₁ lines and controls. Red clusters (A02 SNPs) and blue clusters (A09 SNPs) comprise genotypes without the *A. stenosperma*-derived allele. *A. stenosperma* was genotyped with the alternate allele. Yellow

clusters indicate the backcross lines with incorporated PRKN resistance from *A. stenosperma*.

The principal component analysis (PCA) performed on the genotyping data, allowed us to observe that the BC₃F₁s have recovered a high percentage of the *A. hypogaea* genome. Backcross lines showed a recurrent parent genome recovery between 80.2 and 98.8%, while still carrying between 1.1 and 19.1% of the wild donor genome (Supplementary Figure 2). Additionally, to visualize the distribution of the backcross population according to the proportion of wild introgression (%) in each chromosome, the data was displayed in violin plots. These plots allowed us to observe that the lines were harboring more wild alleles in the A-subgenome than in B-subgenome, especially in chromosomes A02 and A09 where foreground selection was applied during the backcrossing process (Figure 3 and Supplementary Table 7).

Peanut Root-Knot Nematode Resistance Validation Using BC₃F₃s

To validate the PRKN resistance controlled by QTL in A02 and/or A09, 72 BC₃F₃ segregating lines, and resistant and susceptible controls were evaluated using a greenhouse pot nematode bioassay and genotyped with both KASP and Affymetrix to confirm the presence of the QTL. Specifically, Affymetrix was completed for genome-wide characterization. Gall index (GI) and number of eggs in relation to root weight (GI/g and eggs/g) allowed us to assess the resistance to *M. arenaria* within the backcross lines (Table 1).

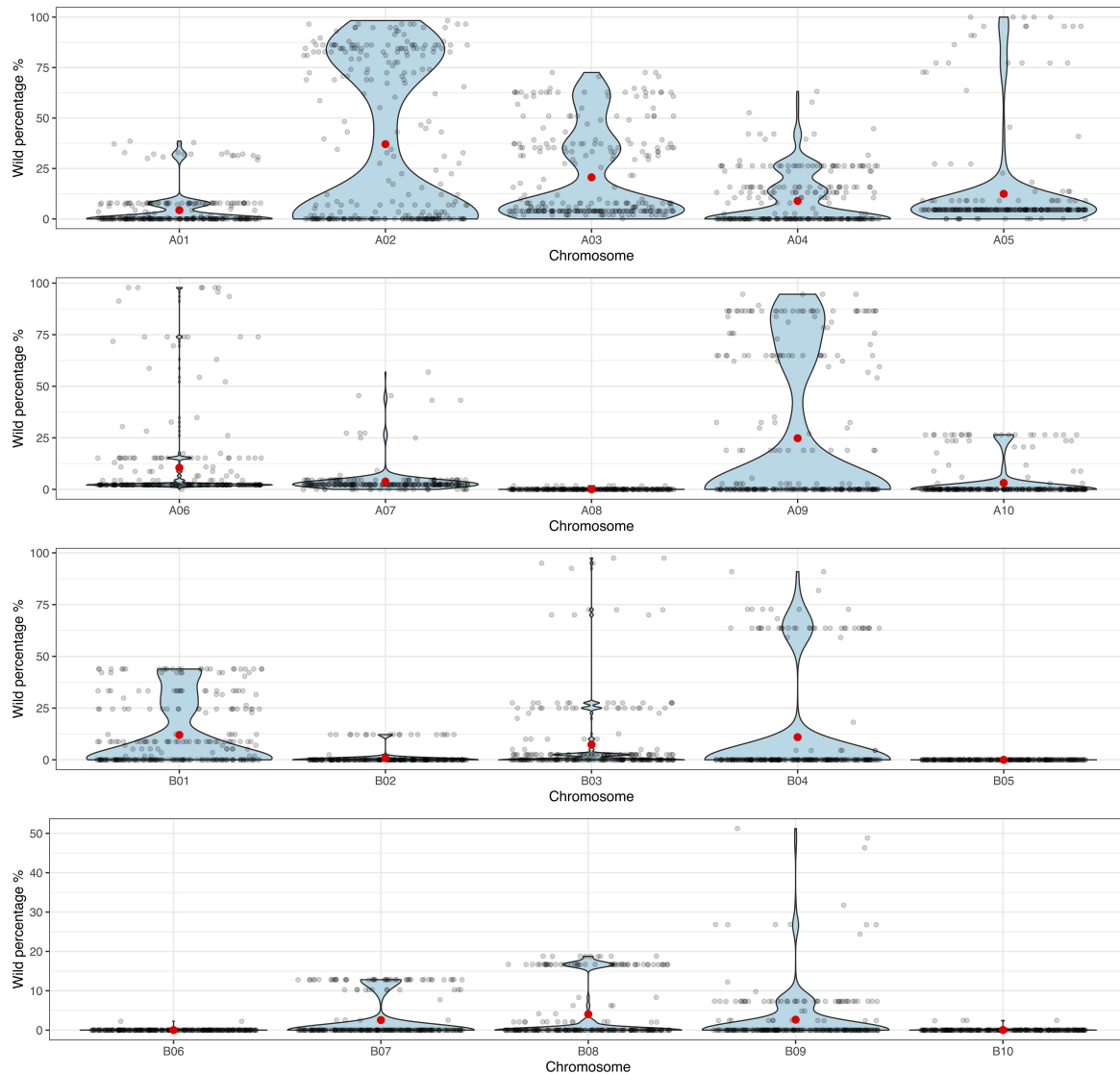


FIGURE 3 | Violin plots for the proportion of wild genome (%) (y-axis) in each of the 10 A- and 10 B-subgenome chromosomes (x-axis) for the 271 BC₃F₁ lines. Black dots indicate each individual BC₃F₁, and red dots indicate the mean.

Resistant controls (BatSten1 and TifNV-High O/L) and susceptible genotypes (5-646-10 and 13-1014) exhibited the expected phenotype. BatSten1 and TifNV-High O/L showed strong resistance, with no or low gall/egg production. In contrast, for 5-646-10 and 13-1014, GI/g fluctuated between 0.31 ± 0.26 and 0.42 ± 0.11 and eggs/g varied between 579.18 ± 855.26 and 1268.64 ± 1046.49 (Figure 4, Table 1 and Supplementary Table 8). BC₃F₃ lines showed significant differences for GI/g and eggs/g (Kruskal-Wallis, Skillings-Mack and Wilcoxon tests, $P < 0.05$). Since the BC₃F₃s were still recombining and segregating for PRKN QTL, to better summarize the results, the lines were grouped according to the segments they were carrying as follows: ● Group 1: bottom A02 (A02) (*A. stenosperra* allele at A02-83,464,195, A02-92,077,207 and A02-92,983,792 → 81.0 – 93.8 Mb); ● Group

2: bottom A09 (A09) (*A. stenosperra* allele at A09-112,309,231 and A09-114,515,959 → 104.6 – 119.8 Mb); ● Group 3: large A09 (A09+) (*A. stenosperra* allele at A09-16,516,448, A09-112,309,231 and A09-114,515,959 → 3.4 – 118.7 Mb); and ● Group 4: Both bottom small A02 (*A. stenosperra* allele at A02-92,077,207 and A02-92,983,792 → 91.6 – 93.8 Mb) and bottom A09 (A02- and A09). According to this grouping, all the backcross materials belonging to groups 1, 3, and 4 exhibited high levels of resistance to PRKN. No or few galls (0.01 ± 0.03 – 0.17 ± 0.20) and low egg production (0.00 ± 0.00 – 36.30 ± 83.17) was observed in the infected roots. For groups 3 and 4 galls were observed in the roots, but the production of eggs was inhibited. In contrast, lines in group 2 were susceptible to PRKN by having GI/g and eggs/g values of 0.26 ± 0.19 and 1050.62 ± 1005.38 , respectively (Figure 4). For more details of

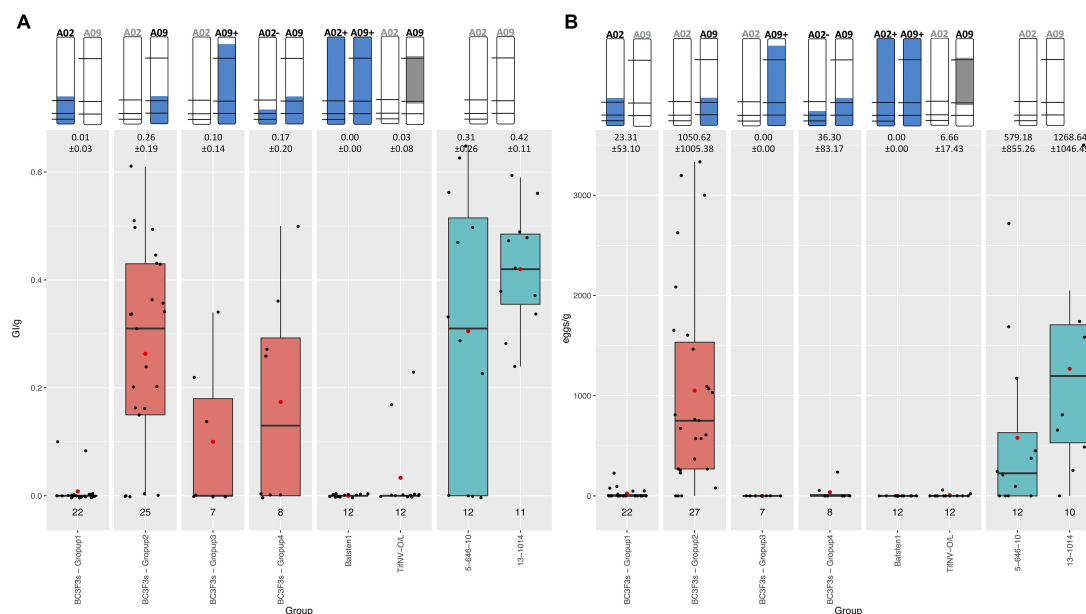


FIGURE 4 | Boxplot diagrams for Galling index per gram of root (GI/g) **(A)** and Number of eggs per gram of root (eggs/g) **(B)** of BC₃F₃ lines, resistant controls Batsten1 and *A. hypogaea* TifNV-High O/L, and susceptible controls 5-646-10 and 13-1014. BC₃F₃ lines were grouped according to the *A. stenospemma* alleles that they are carrying as: ● Group 1: bottom A02 (A02) (*A. stenospemma* allele at A02-83,464,195, A02-92,077,207, and A02-92,983,792); ● Group 2: bottom A09 (A09) (*A. stenospemma* allele at A09-112,309,231 and A09-114,515,959); ● Group 3: large A09 (A09+) (*A. stenospemma* allele at A09-16,516,448, A09-112,309,231 and A09-114,515,959); and ● Group 4: Both bottom small A02 (*A. stenospemma* allele at A02-92,077,207 and A02-92,983,792) and bottom A09 (A02- and A09). The top of the figure indicates the introgression in A02 and A09 from *A. stenospemma* in blue and SNP markers as black horizontal lines. For TifNV-High O/L, resistance from *A. cardenasii* is colored in gray. The numbers at the top of the bars indicate the mean \pm SD. The numbers above the groups indicate the number of lines included in each group. Complete pedigree in **Supplementary Table 8**. Black bars across boxes indicate the median and red dot the mean. BC₃F₃s in salmon color and controls in teal color.

GI/g and eggs/g values for the BC₃F₃ lines, see **Table 1** and **Supplementary Table 8**.

Phenotypic Characterization

A wide variation of morphological and agronomic traits was observed in the backcross populations (BC₃F₁s and BC₃F₂s), including seed size, pollen viability, leaf spot incidence, fertility (number of pegs), plant architecture, flower color, branching, and extra leaves (**Figure 2**). Association analysis between phenotypic information and wild introgressions in the BC₃F₁ population was performed to identify candidate regions associated with these traits.

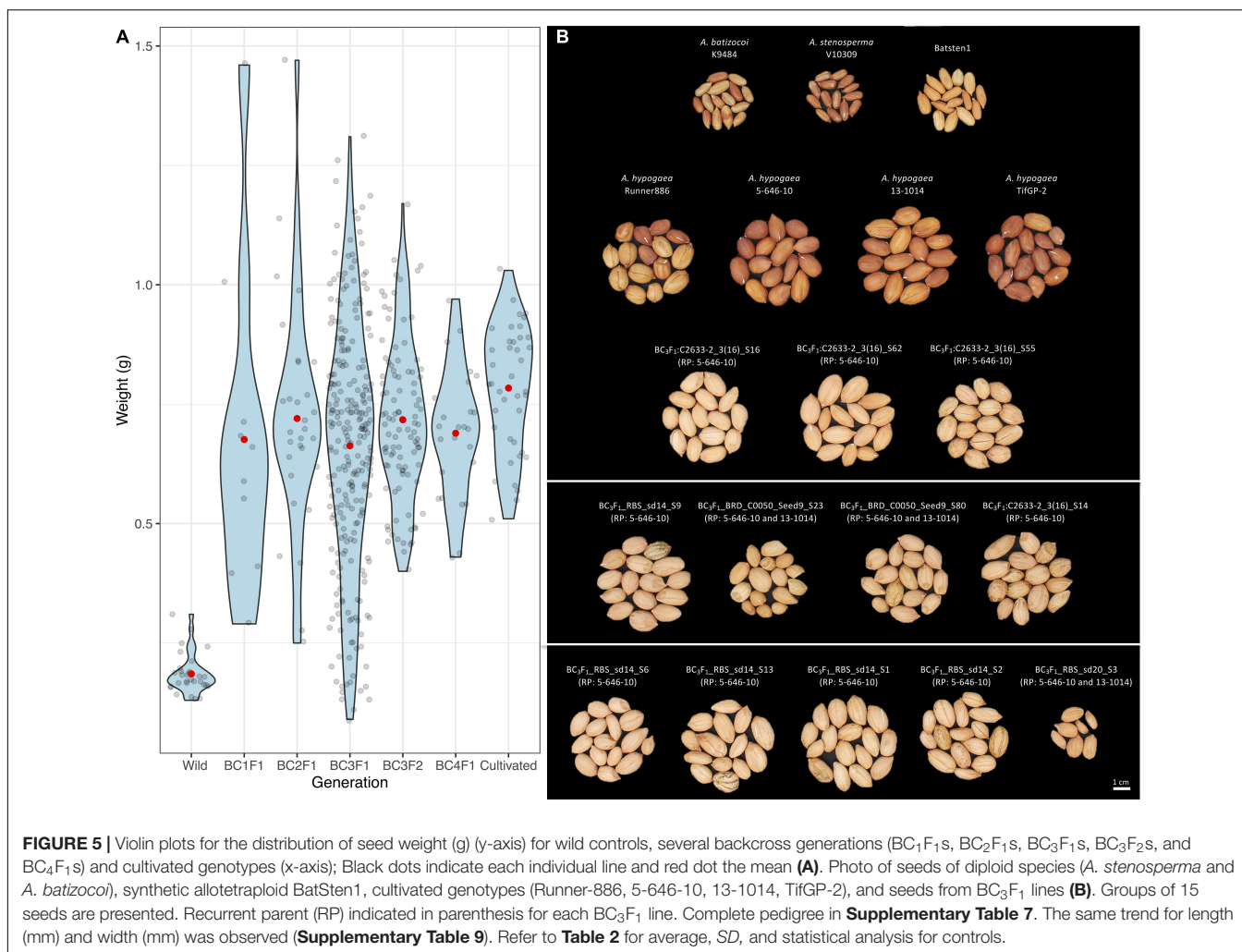
Seed Size

Seed weight (g), length (mm) and width (mm) measurements were recorded for 11 BC₁F₁, 30 BC₂F₁, 253 BC₃F₁, 101 BC₃F₂ and 25 BC₄F₁ seeds prior to planting, along with wild and cultivated controls. Significant differences were observed between the control genotypes according to the Kruskal-Wallis test and the Wilcoxon Test ($P < 0.05$), where wild genotypes have significantly smaller and lighter seeds as compared with seed dimensions of cultivated genotypes (**Figure 5A**, **Table 2**, and **Supplementary Table 9**). There was a clear recovery in seed size as early as BC₁. According to the Welch t -test, the wild accessions (*A. stenospemma* V10309, *A. batizocoi* K9484, and BatSten1) differed significantly from the cultivated genotypes (*A. hypogaea*

5-646-10, 13-1014, TifGP-2 and Runner-886) and the backcross generations (BC₁F₁, BC₂F₁, BC₃F₁, BC₃F₂, and BC₄F₁) for weight, length, and width ($P < 0.05$). When comparing seed measurements of cultivated genotypes with each of the BC generations and between BC generations, in most of the cases there were no significant differences in seed size ($P < 0.05$ and $P < 0.01$) (Welch t -test matrix in **Supplementary Table 9**). Between the BC lines seed size exhibited variation, and on average the seed weight fluctuated from 0.66 g (BC₃F₁s) and 0.72 g (BC₂F₁s and BC₃F₂s), but having weight as high as 1.47 g when compared to the cultivated controls that exhibited similar seed weight of 0.74 g (Runner-886) and 0.84 g (13-1014), with the highest being 1.03 g. In this study, we presented data for the weight (**Figure 5A**), a similar tendency was observed for length and width (**Table 2** and **Supplementary Table 9**). Based on the Pearson correlations performed between the seed weight, length, and width and the genotypic data for the BC₃F₁s potential *A. hypogaea* loci associated with large seed size were identified in chromosomes A03, B01, and B08 (**Supplementary Figures 3A–C** and **Supplementary Table 10**).

Pollen Viability (BC₂F₁s)

Pollen viability (PV) of wild accessions *A. stenospemma*, *A. batizocoi*, the induced allotetraploid BatSten1, recurrent parents 5-646-10 and 13-1014 and 9 BC₂F₁ lines was evaluated to estimate differences in viability within BC lines. Individuals



showed varying degrees of pollen viability, ranging from 65.6 to 94.9% (**Table 3**). In control genotypes, PV varied from 71.7 to 94.9% and in the BC₂F₁s, it fluctuated from 65.6 to 89.8%. Although significant differences were observed between the genotypes according to the Kruskal-Wallis test and the Wilcoxon Test ($P < 0.05$), no grouping trend was observed. In summary, high pollen viability in the cultivated genotypes (5-646-10 and 13-1014), *A. stenosperma*, and some BC₂F₁s were observed, and lower pollen viability was observed for BatSten1, *A. batizocoi*, and other BC₂F₁s (**Supplementary Figure 4**). A low correlation between pollen viability and the number of produced pods per plant was observed (-0.007 , $p < 0.05$).

Leaf Spot Incidence, Fertility, Architecture, and Flower Color (BC₃F₁s)

For the BC₃F₁ population that included 253 lines, segregation for leaf spot incidence, fertility, architecture, and flower color was noticed while growing in the greenhouse in a randomized position (**Figures 2A–C**). 150 lines exhibited signs of leaf spot, while 103 did not (**Supplementary Figure 5A**). Pearson correlation performed between phenotypic and genotypic data allowed us to identify two candidate loci in chromosomes A06

and B02 that could be associated with resistance to leaf spot (**Supplementary Figure 3D**) and are now being tested using BC₃F₂ progeny (data not shown). The number of pegs was counted as an indication of fertility. The distribution of the data showed that the majority of the lines had a similar number of pegs as the recurrent parents 5-646-10 and 13-1014, with an average of 13.15 ± 7.83 total number of pegs (**Supplementary Figure 5B**). For this trait, *A. hypogaea* loci in chromosomes A02 and A09 could be associated with fewer pegs (**Supplementary Figure 3E**). For architecture, around two-thirds of the lines had the cultivated growth habit phenotype (163) and one-third of the lines were exhibiting an erect growth habit (84) (six were too small to be scored) (**Figure 2B** and **Supplementary Figure 5C**). According to the association analysis, SNPs in chromosomes A01 and B08 could be associated with changes in plant architecture (**Supplementary Figure 3F**). Most lines exhibited orange flowers (245) while eight had the yellow color, a common trait within *Arachis* wild species (**Supplementary Figure 5D**). After running the mixed linear model in Tassel, introgression on the top of chromosome A05 (7.92–8.62 Mb) was found to be associated with the change in flower color from orange to yellow. In this study, four markers were found to be linked with this trait

(A05-7,919,003, A05-7,958,564, A05-8,040,921, A05-8,621,849) (**Supplementary Figure 6**).

Leaf Spot Incidence, Architecture, Branching, and Extra Leaves (BC₃F₂s)

Segregation for leaf spot incidence, architecture, branching, and extra leaves was also noticed among the 101 BC₃F₂ lines (two of them were too small to be scored) (**Supplementary Table 11**). 38 lines had leaf spots and 61 did not (**Supplementary Figure 5E**). Around half of the lines had the cultivated growth habit phenotype (51), some exhibiting an erect phenotype (41) and a few with a dwarf phenotype (7) (**Supplementary Figure 5F**). Most lines showed the peanut phenotype of normal branching (79 with score 3) and some presented a high number of stems (20 with score 1 or 2) (**Supplementary Figure 5G**). Finally, at least one extra leaflet was observed in 70 plants (score 0) and 31 did not display this unexpected phenotype (score 1) (**Supplementary Figure 5H**).

DISCUSSION

Crop wild relatives (CWR) have become an important source to reintroduce genetic diversity for crop improvement (Dempewolf et al., 2017). For peanut breeding, diploid wild relatives comprise a diverse genetic pool that is being used to broaden peanut's genetic base (Stalker, 2017). Transferring wild beneficial alleles requires an additional step of developing peanut compatible wild-derived synthetic allotetraploids (Suassuna et al., 2020). To incorporate root-knot nematode resistance from *A. cardenasii*, synthetic allotetraploids were successfully developed and used (Simpson et al., 1993; Simpson and Starr, 2001). The introgression in A09 controlling PRKN resistance (Nagy et al., 2010; Chu et al., 2016) is present in several commercial cultivars (Simpson et al., 2003, 2013; Holbrook et al., 2008, 2017; Branch and Brennenman, 2015) and has provided a strong resistance over the years. Since the resistance to PRKN is controlled by a single source, there is a risk of virulent nematode populations developing. Therefore, the incorporation of new alleles is essential to provide stronger and more durable resistance against PRKN.

One of the peanut wild relatives *A. stenosperrya* PI666100/V10309 has been confirmed to be resistant to peanut root-knot nematode, *M. arenaria* (Proite et al., 2008), and genes involved in plant defense against this pathogen have been described (Proite et al., 2007; Guimarães et al., 2010; Mota et al., 2018; Araujo et al., 2021). The present work reports the successful incorporation of two new and strong PRKN resistance loci from *A. stenosperrya* previously mapped and validated (Leal-Bertioli et al., 2016; Ballén-Taborda et al., 2019, 2021). Here, marker-assisted backcross breeding was employed to complete four cycles, and genetic and phenotypic characterization was performed (**Figure 1**). As pyramiding of major R-genes has been proven to be valuable to extend durability and effectiveness of major genes (Pilet-Nayel et al., 2017), in the fourth cycle the elite breeding line 13-1125 harboring nematode resistance from *A. cardenasii* was included

as a female parent (Holbrook, CC, unpublished data). Based on KASP genotyping, three BC lines could have pyramided both sources of resistance in A02 from *A. stenosperrya* and A09 from *A. cardenasii* (**Supplementary Tables 6, 12**, highlighted in green).

A population of 253 true third backcross lines were subjected to genome-wide genotyping and phenotypically characterized for association analysis. These lines originally selected for PRKN resistance had wild introgressions between 1.1 and 19.1% across the genome. Having a high percentage of the wild genome would indicate that additional cycles of backcrossing are needed to assure maximum elite genome representation; therefore, we completed the fourth cycle. However, segregation for several phenotypic attributes (seed size, pollen viability, leaf spot incidence, fertility, architecture, flower color, branching, and extra leaves) was observed, which indicated that wild introgressions in different chromosomes could be controlling these traits and are worth further study. The whole BC₃ population has been carried forward not only to develop nematode-resistant cultivars but additionally to study resistances to other pests and diseases and to develop a CSSL-like population that would be useful for precise mapping of QTL.

Based on the data of agronomic and morphological traits measured in several generations, most lines exhibited elite peanut traits and were similar to recurrent parental breeding lines. First, seed size in each generation of introgression was examined. This was done to observe the seed size recovery as we progressed through our MABC scheme, as seed size is an important trait associated with germination, vigor, and yield, and is important for the peanut industry and market (Singh et al., 1998). For the backcross lines, our measurements showed that seed size was not progressively increased as the wild genome representation was reduced in each generation. On the contrary, we observed that only one backcross generation was required to recover the elite peanut's large seeds and that in later generations the average seed size did not change significantly (**Figure 5B**). We also observed that larger and heavier seeds were produced by some BC lines, by having seed weight as high as 1.47 g compared to the cultivated controls that exhibited 1.03 g as the heaviest seed. Similar behavior was observed for length and width. An explanation for the transgressive segregation in seed size could be due to the presence of wild alleles that are contributing to larger seed size, as reported before in interspecific peanut progenies (Fonceka et al., 2012a; Suassuna et al., 2015). Although the correlation analysis showed candidate regions in chromosomes A03, B01, and B08, this requires further validation as these have not been reported previously.

Results for pollen viability measured in the BC₂ generation agreed with that of Leal-Bertioli et al. (2015) as cultivated peanut genotypes and *A. stenosperrya* showed a high pollen viability (average of 91.9%) and most of the BC₂F₁ lines had lower numbers (average of 79.2%) reflecting the genetic distance of the parental genotypes. Our results showed little variation regardless of the genotype, which suggests that pollen viability is not a key contributor to the production of seeds

within the backcross lines (Leal-Bertioli et al., 2015). This was also corroborated by the low correlation between pollen staining and pod production (-0.007 , $P < 0.05$). The correlation analysis on the BC₃ generation, also allowed us to identify candidate introgressions for leaf spot reduction in chromosomes A06 and B02. For flower color, the region at the top of chromosome A05 (7.92 – 8.62 Mbp) associated with the yellow color trait is consistent with previous reports, and a result of homologous recombination (Fonceka et al., 2012b; Bertioli et al., 2019).

Although most of the BC lines had domesticated features, we also observed some variation in fertility (number of pegs), plant architecture, branching, and extra leaves. Further analyses will be required to fully understand the candidate wild introgressions that are controlling these traits. Finally, in the case of extra leaflets, this is a phenotype that has been described as a novel heterozygous trait that continues to segregate even after several generations of selfing (Branch et al., 2020).

Validation of Nematode Resistance

In this study, resistance to PRKN was successfully validated in a set of BC₃F₃ lines. The bottom of A02 (81.0 – 93.8 Mb) (Figure 4, group 1) provided strong resistance as previously described and validated (Leal-Bertioli et al., 2016; Ballén-Taborda et al., 2019, 2021). In the case of the QTL in chromosome A09, we observed that the small introgression at the bottom (104.6 – 119.8 Mb) was insufficient to stop nematode development (group 2) and that a larger segment at the top-bottom of A09 (3.4 – 118.7 Mb) was required to provide resistance, especially for preventing eggs production (group 3). It is possible that the presence of wild alleles in A04 associated with susceptibility in group 2 could be acting against resistance as previously described (Ballén-Taborda et al., 2019). When the plants were harboring both bottom small A02 (91.6 – 93.8 Mb) and bottom A09 (104.6 – 119.8 Mb) (group 4) we would expect to observe inhibition of both galls and egg production since A02 was present; however, galls were present in the roots. Field testing is in progress to test the stability of the resistance and for allele fixation through selfing.

Implications for Breeding for Disease Resistance

Genetic maps, quantitative trait loci (QTL), and marker-phenotype associations have been reported for numerous crops and traits (Collard and Mackill, 2008). Despite this, examples of QTL incorporation in plant breeding programs are lower than expected (Bernardo, 2016). This work represents a successful example of QTL introgression from a wild relative into an elite peanut despite the genetic incompatibilities. This provides an alternative to the only source of root-knot resistance currently deployed in the peanut crop, derived from *A. cardenasii* (Simpson and Starr, 2001).

The population of advanced peanut backcross lines that we have developed during this work has wild chromosome segments through much of the genome, distributed in different ways in different lines. They are being tested and advanced in several

locations, and the best performing lines are being selected for germplasm release. The PRKN resistance alleles have been successfully validated and DNA markers are now available to facilitate the marker-assisted selection. Furthermore, because of the diverse wild chromosome segments in this population, we also anticipate that it has other disease resistances and traits of value to the peanut crop. We anticipate that, over time, these backcrossed lines will impact peanut production by delivering several new traits to the peanut crop, similar to the case of North Carolina peanut lines with *A. cardenasii* segments that have provided resistance to late leaf spot, rust, and web blotch in numerous countries around the world (Bertioli et al., 2021b).

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.

AUTHOR CONTRIBUTIONS

SCML-B, DJB, SAJ, and PO-A conceptualized this project. PO-A, YC, CCH, and CB-T performed crosses. CB-T and PT performed phenotyping for nematode resistance. CB-T performed phenotypic characterization of traits and wrote the original draft. CB-T, DJB, and SCML-B conducted genotyping and data analysis. All authors revised and approved the manuscript.

FUNDING

This work was supported by the Agriculture and Food Research Initiative Competitive Grant (#2018-67013-28139) co-funded by the USDA National Institute of Food and Agriculture and the National Peanut Board and grants from the National Science Foundation (#MCB-1543922), Georgia Peanut Commission, Peanut Research Foundation, and Mars Wrigley Inc.

ACKNOWLEDGMENTS

We appreciate Jenny Leverett, Victoria Morris, Mark Hopkins, Stephanie Botton, and Emily Jackson for their technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.785358/full#supplementary-material>

REFERENCES

- Araujo, A. C. G., Guimarães, P. M., Mota, A. P. Z., Guimarães, L. A., Pereira, B. M., Vinson, C. C., et al. (2021). Overexpression of DUF538 from wild *arachis* enhances plant resistance to *Meloidogyne* spp. *Agronomy* 11:559. doi: 10.3390/agronomy11030559
- Ballén-Taborda, C., Chu, Y., Ozias-Akins, P., Timper, P., Holbrook, C. C., Jackson, S. A., et al. (2019). A new source of root-knot nematode resistance from *Arachis stenosperma* incorporated into allotetraploid peanut (*Arachis hypogaea*). *Sci. Rep.* 9:17702. doi: 10.1038/s41598-019-54183-1
- Ballén-Taborda, C., Chu, Y., Ozias-Akins, P., Timper, P., Jackson, S. A., Bertoli, D. J., et al. (2021). Validation of resistance to root-knot nematode incorporated in peanut from the wild relative *Arachis stenosperma*. *Agron. J.* 113, 2293–2302. doi: 10.1002/agj.20654
- Bernardo, R. (2016). Bandwagons I, too, have known. *Theor. Appl. Genet.* 129, 2323–2332. doi: 10.1007/s00122-016-2772-5
- Bertoli, D. J., Abernathy, B., Seijo, G., Clevenger, J., and Cannon, S. B. (2020). Evaluating two different models of peanut's origin. *Nat. Genet.* 52, 557–559.
- Bertoli, D. J., Cannon, S. B., Froenicke, L., Huang, G., Farmer, A. D., Cannon, E. K., et al. (2016). The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat. Genet.* 48, 438–446. doi: 10.1038/ng.3517
- Bertoli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., et al. (2019). The genome sequence of peanut (*Arachis hypogaea*), a segmental allotetraploid. *Nat. Genet.* 51, 877–884. doi: 10.1038/s41588-019-0405-z
- Bertoli, D. J., Gao, D., Ballén-Taborda, C., Chu, Y., Ozias-Akins, P., Jackson, S., et al. (2021a). Registration of GA-BatSten1 and GA-MagSten1, two induced allotetraploids derived from peanut wild relatives with superior resistance to leaf spots, rust and root-knot nematode. *J. Plant Regist.* 15, 372–378. doi: 10.1002/plr.20133
- Bertoli, D. J., Clevenger, J., Godoy, I. J., Stalker, H. T., Wood, S., Santos, J. F., et al. (2021b). Legacy genetics of *Arachis cardenasii* in the peanut crop shows the profound benefits of international seed exchange. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2104899118. doi: 10.1073/pnas.2104899118
- 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
- Branch, W. D., and Brenneman, T. B. (2015). Registration of 'Georgia-14N' peanut. *J. Plant Regist.* 9, 159–161.
- Branch, W. D., Tallury, S. P., Clevenger, J. P., Schwartz, B. M., and Hanna, W. W. (2020). Inheritance of a novel heterozygous peanut mutant, 5-small leaflet. *Peanut Sci.* 47, 33–37.
- Chatfield, M., and Mander, A. (2009). The skillings-mack test (Friedman test when there are missing data). *Stata J.* 9, 299–305.
- Chu, Y., Gill, R., Clevenger, J., Timper, P., Corley Holbrook, C., and Ozias-Akins, P. (2016). Identification of rare recombinants leads to tightly linked markers for nematode resistance in peanut. *Peanut Sci.* 43, 88–93.
- Clevenger, J. P., Korani, W., Ozias-Akins, P., and Jackson, S. (2018). Haplotype-based genotyping in polyploids. *Front. Plant Sci.* 9:564. doi: 10.3389/fpls.2018.00564
- Collard, B. C., and Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170
- Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., and Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Sci.* 57, 1070–1082.
- Dickson, D. W., and De Waele, D. (2005). "Nematode parasites of peanut," in *Plant Parasitic Nematodes in Subtropical and Tropical Agriculture*, eds R. A. Sikora and J. Bridge (Wallingford: CABI Publishing), 393–436. doi: 10.1079/9780851997278.0393
- FAOSTAT (2021). *FAOSTAT Statistical Database*. Rome: Food and Agriculture Organization of the United Nations.
- Fonckea, D., Tossim, H. A., Rivallan, R., Vignes, H., Faye, I., Ndoye, O., et al. (2012a). Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. *BMC Plant Biol.* 12:26. doi: 10.1186/1471-2229-12-26
- Fonckea, D., Tossim, H.-A., Rivallan, R., Vignes, H., Lacut, E., De Bellis, F., et al. (2012b). Construction of chromosome segment substitution lines in peanut (*Arachis hypogaea* L.) using a wild synthetic and QTL mapping for plant morphology. *PLoS One* 7:e48642. doi: 10.1371/journal.pone.0048642
- Gaaliche, B., Majdoub, A., Trad, M., and Mars, M. (2013). Assessment of pollen viability, germination, and tube growth in eight tunisian *Caprifig* (*Ficus carica* L.) Cultivars. *ISRN Agron.* 2013, 1–4.
- Guimarães, P., Brasileiro, A. C. M., Proite, K., Araujo, A., Leal-Bertoli, S., Pic-Taylor, A., et al. (2010). A study of gene expression in the nematode resistant wild peanut relative, *Arachis stenosperma*, in response to challenge with *Meloidogyne arenaria*. *Trop. Plant Biol.* 3, 183–192.
- Hamazaki, K., and Iwata, H. (2020). RAINBOW: haplotype-based genome-wide association study using a novel SNP-set method. *PLoS Comput. Biol.* 16:e1007663. doi: 10.1371/journal.pcbi.1007663
- Heslop-Harrison, J. S. (1992). "Cytological techniques to assess pollen quality," in *Sexual Plant Reproduction*, eds M. Cresti and A. Tiezzi (Berlin: Springer-Verlag), 41–48. doi: 10.1007/978-3-642-77677-9_4
- Holbrook, C., and Stalker, H. (2003). Peanut breeding and genetic resources. *Plant Breed. Rev.* 22, 297–356. doi: 10.1002/9780470650202.ch6
- Holbrook, C., Dong, W., Timper, P., Culbreath, A., and Kvien, C. (2012). Registration of peanut germplasm line TifGP-2, a nematode-susceptible sister line of 'Tifguard'. *J. Plant Regist.* 6, 208–211. doi: 10.3198/jpr2011.09.0507crg
- Holbrook, C., Ozias-Akins, P., Chu, Y., Culbreath, A. K., Kvien, C., and Brenneman, T. (2017). Registration of 'TifNV-High O/L' peanut. *J. Plant Regist.* 11, 228–230.
- Holbrook, C., Timper, P., and Culbreath, A. (2003). Resistance to and root-knot nematode in peanut interspecific breeding lines. *Crop. Sci.* 43, 1109–1113.
- Holbrook, C., Timper, P., Culbreath, A., and Kvien, C. (2008). Registration of 'Tifguard' peanut. *J. Plant Regist.* 2, 391–410.
- Hussey, R. S., and Barker, K. R. (1973). A comparison of methods of collecting inocula of *Meloidogyne* species, including a new technique. *Plant Dis. Report.* 57, 1025–1028.
- Korani, W., Clevenger, J. P., Chu, Y., and Ozias-Akins, P. (2019). Machine learning as an effective method for identifying true single nucleotide polymorphisms in polyploid plants. *Plant Genome* 12:180023. doi: 10.3835/plantgenome2018.05.0023
- Kruskal, W., and Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* 47, 583–621. doi: 10.1080/01621459.1952.10483441
- Leal-Bertoli, S. C. M., Moretzsohn, M. C., Roberts, P. A., Ballén-Taborda, C., Borba, T. C. O., Valdisser, P. A., et al. (2016). Genetic mapping of resistance to *Meloidogyne arenaria* in *Arachis stenosperma*: a new source of nematode resistance for peanut. *G3* 6, 377–390. doi: 10.1534/g3.115.023044
- Leal-Bertoli, S. C., Santos, S. P., Dantas, K. M., Inglis, P. W., Nielsen, S., Araujo, A. C., et al. (2015). *Arachis batizocoi*: a study of its relationship to cultivated peanut (*A. hypogaea*) and its potential for introgression of wild genes into the peanut crop using induced allotetraploids. *Ann. Bot.* 115, 237–249. doi: 10.1093/aob/mcu237
- Mota, A. P. Z., Vidigal, B., Danchin, E. G. J., Togawa, R. C., Leal-Bertoli, S. C. M., Bertoli, D. J., et al. (2018). Comparative root transcriptome of wild *Arachis* reveals NBS-LRR genes related to nematode resistance. *BMC Plant Biol.* 18:159. doi: 10.1186/s12870-018-1373-7
- Nagy, E. D., Chu, Y., Guo, Y., Khanal, S., Tang, S., Li, Y., et al. (2010). Recombination is suppressed in an alien introgression in peanut harboring Rma, a dominant root-knot nematode resistance gene. *Mol. Breed.* 26, 357–370. doi: 10.1007/s11032-010-9430-4
- Nelson, S. C., Simpson, C. E., and Starr, J. L. (1989). Resistance to *Meloidogyne arenaria* in *Arachis* spp. germplasm. *J. Nematol.* 21, 654–660.
- Noe, J. P., Holbrook, C. C., and Minton, N. A. (1992). Field evaluation of susceptibility to *Meloidogyne arenaria* in *Arachis hypogaea* plant introductions. *J. Nematol.* 24, 712–716.
- Oka, Y. (2020). From old-generation to next-generation nematocides. *Agronomy* 10:1387. doi: 10.3390/agronomy10091387
- Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., et al. (2017). Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Front. Plant Sci.* 8:1838. doi: 10.3389/fpls.2017.01838

- Pittman, R. N. (1995). *United States Peanut Descriptors*. U.S. Department of Agriculture, Agricultural Research Service, ARS-132. Tifton, GA: U.S. Department of Agriculture, Agricultural Research Service, 18.
- Proite, K., Carneiro, R., Falcão, R., Gomes, A., Leal-Bertioli, S., Guimarães, P., et al. (2008). Post-infection development and histopathology of *Meloidogyne arenaria* race 1 on *Arachis* spp. *Plant Pathol.* 57, 974–980.
- Proite, K., Leal-Bertioli, S. C., Bertioli, D. J., Moretzsohn, M. C., Da Silva, F. R., Martins, N. F., et al. (2007). ESTs from a wild *Arachis* species for gene discovery and marker development. *BMC Plant Biol.* 7:7. doi: 10.1186/1471-2229-7-7
- Simpson, C., and Starr, J. (2001). Registration of ‘COAN’ peanut. *Crop Sci.* 41:918. doi: 10.2135/cropsci2001.413918x
- Simpson, C., Nelson, S. C., Starr, J. L., Woodard, K. E., and Smith, O. D. (1993). Registration of TxAG-6 and TxAG-7 peanut germplasm. *Crop Sci.* 33:1418.
- Simpson, C., Starr, J. L., Baring, M. R., Burow, M. D., Cason, J. M., and Wilson, J. N. (2013). Registration of ‘Webb’ peanut. *J. Plant Regist.* 7:265. doi: 10.1111/cea.13816
- Simpson, C., Starr, J. L., Church, G., Burow, M. D., and Paterson, A. H. (2003). Registration of NemaTAM peanut. *Crop Sci.* 43:1561. doi: 10.2135/cropsci2003.1561
- Singh, A., Nautiyal, P., and Zala, P. V. (1998). Growth and yield of groundnut varieties as influenced by seed size. *Trop Sci.* 38, 48–56. doi: 10.1186/s13002-018-0275-y
- Stalker, T. H. (2017). Utilizing wild species for peanut improvement. *Crop Sci.* 57:1102. doi: 10.2135/cropsci2016.09.0824
- Suassuna, T. M. F., Suassuna, N. D., Moretzsohn, M. C., Leal-Bertioli, S. C. M., Bertioli, D. J., and Medeiros, E. P. (2015). Yield, market quality, and leaf spots partial resistance of interspecific peanut progenies. *Crop Breed. Appl. Biotechnol.* 15, 175–180. doi: 10.1590/1984-70332015v15n3n30
- Suassuna, T., Suassuna, N., Martins, K., Matos, R., Heuert, J., Bertioli, D., et al. (2020). Broadening the variability for peanut breeding with a wild species-derived induced allotetraploid. *Agronomy* 10:1917. doi: 10.3390/agronomy10121917
- Taylor, A. L., and Sasser, J. N. (1978). *Biology, Identification and Control of Root-Knot Nematodes (Meloidogyne spp.)*. Raleigh: North Carolina State University, 111–112.
- Timper, P., Dickson, D. W., and Steenkamp, S. (2018). “Nematode parasites of groundnut,” in *Plant Parasitic Nematodes in Subtropical and Tropical Agriculture*, eds R. A. Sikora and D. L. Coyne (Wallingford: CAB International), 411–445.
- Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv [Preprint]* doi: 10.1101/005165
- You, F. M., Huo, N., Gu, Y. Q., Luo, M. C., Ma, Y., Hane, D., et al. (2008). BatchPrimer3: a high throughput web application for PCR and sequencing primer design. *BMC Bioinform.* 9:253. doi: 10.1186/1471-2105-9-253

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 Ballén-Taborda, Chu, Ozias-Akins, Holbrook, Timper, Jackson, Bertioli and Leal-Bertioli. 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 Quantitative Trait Loci Associated With Partial Resistance to Fusarium Root Rot and Wilt Caused by *Fusarium graminearum* in Field Pea

Longfei Wu, Rudolph Fredua-Agyeman, Stephen E. Strelkov, Kan-Fa Chang and Sheau-Fang Hwang*

Department of Agricultural, Food and Nutritional Science, University of Alberta, Edmonton, AB, Canada

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Sandra E. Branham,
Clemson University, United States
Mahesh Damodhar Mahendrakar,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

Lucia Vieira Hoffmann,
Brazilian Agricultural Research
Corporation (EMBRAPA), Brazil

*Correspondence:

Sheau-Fang Hwang
sheau.fang.hwang@ualberta.ca

Specialty section:

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

Received: 28 September 2021

Accepted: 10 December 2021

Published: 20 January 2022

Citation:

Wu L, Fredua-Agyeman R,
Strelkov SE, Chang K-F and
Hwang S-F (2022) Identification
of Quantitative Trait Loci Associated
With Partial Resistance to *Fusarium*
Root Rot and Wilt Caused by
Fusarium graminearum in Field Pea.
Front. Plant Sci. 12:784593.
doi: 10.3389/fpls.2021.784593

Fusarium root rot, caused by a complex of *Fusarium* spp., is a major disease of field pea (*Pisum sativum*). The development of genetic resistance is the most promising approach to manage the disease, but no pea germplasm has been identified that is completely resistant to root rot. The aim of this study was to detect quantitative trait loci (QTL) conferring partial resistance to root rot and wilting, caused by five fungal isolates representing *Fusarium solani*, *F. avenaceum*, *F. acuminatum*, *F. proliferatum*, and *F. graminearum*. Evaluation of the root rot-tolerant cultivar “00-2067” and susceptible cultivar “Reward” was carried out with the five species. There was a significant difference ($p < 0.001$) between the mean root rot values of the two cultivars inoculated with the *F. avenaceum* (F4A) and *F. graminearum* (FG2) isolates. Therefore, in the QTL study, the F₈ recombinant inbred line (RIL) population derived from “Reward” × “00-2067” was inoculated in the greenhouse (4 ×) with only F4A and FG2. The parents and F₈ population were genotyped using 13.2K single nucleotide polymorphisms (SNPs) and 222 simple sequence repeat (SSR) markers. A significant genotypic effect ($p < 0.05$) and high heritability (79% to 92.1%) were observed for disease severity, vigor, and plant height following inoculation with F4A and FG2. Significant correlation coefficients were detected among and within all traits. This suggested that a high proportion of the genetic variance was transmitted from the parents to the progeny. However, no significant QTL (LOD > 3) were detected for the RILs inoculated with F4A. In the case of the RILs inoculated with FG2, 5 QTL for root rot severity and 3 QTL each for vigor and plant height were detected. The most stable QTL for plant height (*Hgt-Ps3.1*) was detected on Chrom5/LGIII. The two most stable QTL for partial resistance to FG2, *Fg-Ps4.1*, and *Fg-Ps4.2* were located in a 15.1-cM and 11.2-cM genomic region, respectively, on Chrom4/LGIV. The most stable QTL for vigor (*Vig-Ps4.1*) was found in the same region. Twenty-five major and moderate effect digenic epistatic interactions were detected. The identified region on chrom4/LGIV could be important for resistance breeding and marker development.

Keywords: *Pisum sativum* L., recombinant inbred lines (RIL), conidia suspension, SNP and SSR markers, linkage map construction and QTL mapping

INTRODUCTION

Globally, root rot is estimated to cause yield reductions of 10–30% in pulse crops, but losses can be as high as 100% in crops with severe infections under ideal environmental conditions (Oyarzun, 1993; Schneider et al., 2001; Schwartz et al., 2005; Cichy et al., 2007). As such, root rot is one of the most devastating diseases of field pea and other pulse crops in Canada and worldwide (Hwang and Chang, 1989; Feng et al., 2010; Chatterton et al., 2015, 2019; Gossen et al., 2016; Chang et al., 2017; Safarieskandari et al., 2020; Wu et al., 2021). The causal organisms of the pea root rot complex (PRRC) are soil-borne fungal and fungal-like organisms that include *Fusarium* spp., *Aphanomyces euteiches*, *Pythium* spp., *Phytophthora* spp., *Rhizoctonia* spp., *Didymella* spp. (formerly *Mycosphaerella* spp.), and *Ascochyta* spp. (Fletcher et al., 1991; Kaiser, 1992; Hwang et al., 1994; Xue et al., 1998; Bailey et al., 2003; Chang et al., 2005, 2013, 2014, 2017; Tyler, 2007; Díaz Arias et al., 2011).

Given their abundance and wide host range, the vast majority of the PRRC organisms are *Fusarium* species, although these may exhibit variable virulence toward different hosts. The various species identified in the Canadian prairies include *F. solani*, *F. avenaceum*, *F. oxysporum*, *F. graminearum*, *F. culmorum*, *F. acuminatum*, *F. redolens*, *F. sambucinum* var. *coeruleum*, *F. equiseti*, *F. poae*, *F. sporotrichioides*, and *F. tabacinum* (Kraft and Pflieger, 2001; Fernandez, 2007; Fernandez et al., 2008; Feng et al., 2010; Chittem et al., 2015; Wu et al., 2017; Chang et al., 2018; Zitnick-Anderson et al., 2018). Among these, *F. avenaceum*, *F. solani*, and *F. oxysporum* were reported to be the primary species causing significant *Fusarium* root rot (FRR) in the major field pea cultivation regions in Canada and worldwide (Kraft, 1981; Kraft and Pflieger, 2001; Wille et al., 2020).

The *Fusarium graminearum* species complex (FGSC) includes the major pathogens causing *Fusarium* head blight (FHB) of wheat, barley, oats, and other small grain cereals (O'Donnell et al., 2008). On cereal hosts, FGSC produces various mycotoxins known as trichothecenes [e.g., deoxynivalenol (DON), nivalenol (NIV), zearalenone (ZEN), and fumonisin B1 (FB1)], which are detrimental to human and animal health when ingested (van der Lee et al., 2015). While *F. graminearum* mainly affects cereals, this pathogen has been isolated from field pea in Canada, the USA, and Lithuania (Feng et al., 2010; Chittem et al., 2015; Rasiukevičiūtė et al., 2019). Rasiukevičiūtė et al. (2019) reported that field pea was the non-cereal crop most susceptible to *F. graminearum* compared with faba bean, fodder beet, oilseed rape, potato, and sugar beet.

At present, there are no sources of complete resistance to PRRC in field pea. Furthermore, higher global temperatures and excessive soil moisture associated with climate change have led to the increased incidence and severity of many plant diseases (Chakraborty et al., 2000; Dorrance et al., 2003; Gautam et al., 2013; Elad and Pertot, 2014). While tillage was reported to be beneficial to the soil environment, it did not suppress the development of FRR in field pea (Bailey et al., 1992). Seedling data and depth were reported to affect FRR in lentil (Hwang et al., 2000), but not in field pea (Chang et al., 2013).

Crop rotations longer than 4 years are recommended for the management of root rot, but these are not always practical (Hwang and Chang, 1989; Bainard et al., 2017). Fungicidal seed treatments were reported to increase emergence and reduce root rot severity in the early growth stages of pea (Xue et al., 2000; Wu et al., 2019), with Apron Maxx (fludioxonil, metalaxyl-M and S-isomer), prothioconazole, fluopyram, and penthiopyrad, suppressing FRR in greenhouse and field experiments (Avenot and Michailides, 2010; Chang et al., 2013). However, some fungicides can also affect *Rhizobia*, leading to reductions in nodulation and nitrogen fixation (Chang et al., 2013), and their use is not environmentally friendly.

Genetic resistance offers the most promising way to control FRR and wilt in pea. However, there is no complete resistance to FRR in field pea, and only a few studies have reported QTL associated with partial resistance to this disease (Feng et al., 2011; McPhee et al., 2012; Coyne et al., 2015, 2019). Coyne et al. (2015, 2019) identified the major QTL for partial resistance to *F. solani*, *Fsp-Ps2.1*, to be on LGII (Chromosome 6), while four minor QTL were found on LGIII, IV, VI, and VII (Chromosomes 5, 4, 1, and 7, respectively). These QTL explained 44.4–53.4% of the total variance for resistance (Coyne et al., 2019). McPhee et al. (2012) detected one major QTL on LGIV (Chromosome 4) and two minor QTL on LGIII (Chromosome 5) to be associated with partial resistance to *F. oxysporum* race 2. The major QTL identified by McPhee et al. (2012), *Fnw4.1*, explained 68–80% of the phenotypic variance. Feng et al. (2011) identified one QTL controlling resistance to *F. avenaceum* on LGVII (Chromosome 7) in a rough map generated with 14 SSRs. The QTL identified in most of these studies had very large confidence intervals due to the limited number of markers used. The low marker density makes it difficult to apply the identified markers in pea breeding.

On the Canadian prairies, cereals are grown in tight rotations with canola, while the cultivation of field pea and other pulses is increasing (Bekkering, 2013; Gill, 2018). Boom-and-bust-type cycles of root rot diseases were highly correlated with crop rotation practices (Govaerts et al., 2007; Su et al., 2021). Therefore, the order of cultivation of crops in a rotation is important. The increased incidence and severity of FRR in field pea make the study of the genetic resistance to different *Fusarium* spp. an important research objective.

Therefore, the objectives of this study were to: (1) evaluate the partially resistant pea cultivar “00-2067” for resistance to different *Fusarium* spp. recovered from surveys for root rot in Alberta, Canada; (2) map the QTL associated with partial resistance to FRR using a segregating recombinant inbred line (RIL) pea population genotyped by Wu et al. (2021) and the most virulent of the *Fusarium* isolates; and (3) determine the stability of the QTL, accounting for disease severity, vigor, and plant height.

MATERIALS AND METHODS

Plant Materials

One-hundred thirty-five RILs used by Wu et al. (2021) for mapping the QTLs associated with partial resistance to *Aphanomyces* root rot were included in this study. In brief, the

Aphanomyces root rot-resistant pea parent “00-2067” developed by Dr. J. Kraft and V. A. Coffman at the Irrigated Agriculture Research and Extension Center in Prosser, WA, United States (Conner et al., 2013; Wu et al., 2021), was used in genetic crosses with the susceptible parent “Reward” (Bing et al., 2006) to produce F₁ plants, which were then used to develop an F₈ RIL population (Supplementary Figure 1) by the single-seed descent (SSD) method (Brim, 1966).

Fusarium Isolates

Five single-spore isolates (SSI), S4C (*F. solani*), F4A (*F. avenaceum*), F037 (*F. acuminatum*), F039 (*F. proliferatum*), and FG2 (*F. graminearum*), representing the *Fusarium* species most frequently recovered from symptomatic pea plants in root rot surveys in Alberta, were used to screen the parental cultivars “00-2067” and “Reward.” Briefly, to obtain the SSIs, surface-sterilized pieces of root tissue with disease lesions were placed on potato dextrose agar (PDA) and incubated at 25°C for 2–3 days and then transferred to the peptone-pentachloronitrobenzene (PCNB) medium for further selection. Mycelial tips of the fungal isolates were cut from selected colonies under a stereomicroscope (Zeiss Axio Scope A1, Carl Zeiss Canada Ltd., Canada), and the water agar (WA) procedure was used to obtain SSI (Zitnick-Anderson et al., 2020). The species designation of each of the SSIs was confirmed based on morphology and evaluation with the PCR primer sets ITS4/ITS5 and EF-1/EF-2, while isolate virulence was confirmed by fulfilling Koch’s postulates (Feng et al., 2010; Chen et al., 2014; Zhou et al., 2014; Wu et al., 2017; Chang et al., 2018; Zitnick-Anderson et al., 2018).

Inoculum Production

Conidial suspensions of the five isolates were generated following Son et al. (2013). Pure cultures of each *Fusarium* spp. were grown in Petri dishes on PDA under darkness at room temperature for 4–6 weeks. Sterile-distilled water was added to each Petri dish, and the surface of each colony was gently scraped with a sterile inoculation needle to dislodge the spores (and hyphal fragments), with the resulting suspension decanted into 250-ml Erlenmeyer flasks, containing a 100-ml autoclaved CMC medium (1.5% carboxymethyl cellulose, 0.1% yeast extract, 0.05% MgSO₄·7H₂O, 0.1% NH₄NO₃, 0.1% KH₂PO₄, 100-ml H₂O). The flasks were covered in aluminum foil to block light and incubated on a rotary shaker at room temperature for 2 weeks. The suspension was centrifuged to collect conidia. The concentration of conidia was estimated with a hemocytometer and adjusted to a final concentration of 2×10^6 spores ml⁻¹ with sterile-deionized water.

Screening of Recombinant Inbred Line Parents With Five *Fusarium* Species

Plastic cups (9-cm diameter and 10.5-cm depth) were filled with a sterilized potting mixture (Cell-Tech™, Monsanto, Winnipeg, MB, Canada). In the greenhouse tests with each SSI (S4C, F4A, F037, F039, and FG2), the roots of seven 5-day-old seedlings of the partially resistant parent “00-2067” and the susceptible parent “Reward” were immersed in the conidial suspension for

15 min and transplanted into the soilless mixture in a cup. An aliquot (1 ml) of conidial suspension was pipetted onto the roots before they were covered with the potting mix. The plants were kept in a greenhouse at 20–25°C/15–18°C day/night and a 16-h photoperiod with daily watering to maintain the potting mix at saturation conditions conducive for FRR development. Each experiment was repeated two times. After 3 weeks, disease severity was estimated for the parental cultivars on a scale of 0–4, where: 0 = completely healthy; 1 = brown or black spots on the main root; 2 = lesions covering the main root, but the rootlets still healthy; 3 = lesions spread to the entire root system; and 4 = root totally dead (Figures 1a,c).

Disease Assessment of Recombinant Inbred Line Population Under Controlled Conditions

The most virulent of the *Fusarium* isolates was used as inoculum to screen the 135 RIL population and parents under greenhouse conditions. The inoculation and the maintenance of the plants were as described above. The pots were arranged in a randomized complete block design (RCBD) with four replicates. The greenhouse experiment was repeated four times. After 3 weeks, plant height (cm) was measured from the base of the stem to the top leaf. Plant vigor was evaluated as a measure of the wilting severity on a scale of 0–4 (4 = plant completely healthy; 3 = thin stem and short height; 2 = brown lesions on stem and yellowing of leaf tips; 1 = wilting on stems and leaves; 0 = plant completely dead) (Figure 1b). The plants were then carefully uprooted, washed under standing water, and assessed for disease severity as described above.

Statistical Analysis of Phenotypic Data

ANOVA was conducted using R software (R Core Team, 2019) for disease severity, vigor, and plant height in four greenhouse

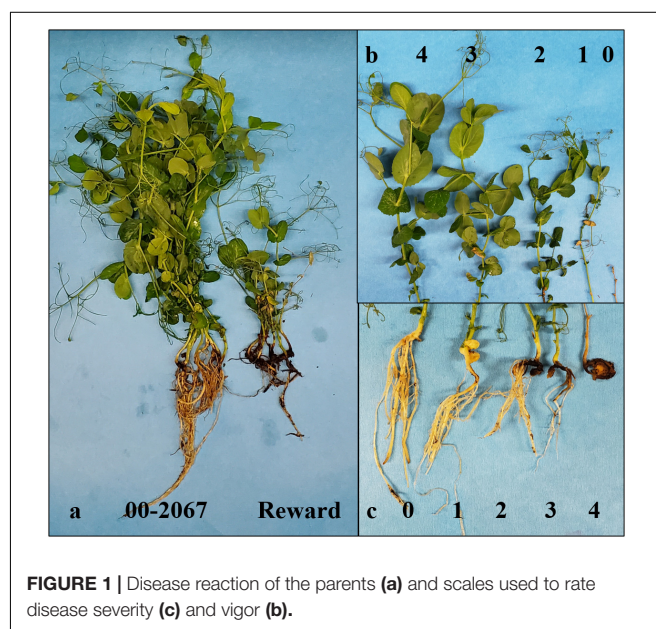


FIGURE 1 | Disease reaction of the parents (a) and scales used to rate disease severity (c) and vigor (b).

environments. The mean and least square mean (LSM) of all traits were calculated for single environments and total data using the package “lsmeans” (Lenth, 2016) in R. To estimate random effects, the best linear unbiased predictors (BLUPs) and heritability were also calculated using the package “Phenotype” (Zhao, 2020) in R. Correlation analysis was conducted within each trait (all variables including means for single environments, LSM, and BLUPs for total data) and among traits (including disease severity, vigor, and plant height) using the package “PerformanceAnalytics” (Peterson et al., 2019) in R, displaying the correlation coefficient, frequency distribution, and dot plot. The *P*-value of the Shapiro–Wilk test was used to determine the normality for each variable using R (R Core Team, 2019).

Genotyping With Single Nucleotide Polymorphisms and Simple Sequence Repeat Markers

The 13.2K SNP markers and 222 SSR markers, the parents, and the RIL population genotyped by Wu et al. (2021) were used in this study. In brief, SNP genotyping was carried out at TraitGenetics GmbH, Gatersleben, Germany, using an SNP array developed from gene-encoding sequences, which are distributed uniformly across the pea genome (Tayeh et al., 2015). The SSR markers were obtained from Loidon et al. (2005). In the case of the SSR markers, the PCR assays, thermal cycling conditions, and genotyping using an ABI PRISM 3730 x1 DNA analyzer (Applied Biosystems, Foster City, CA, United States) were as described by Wu et al. (2021). Filtering of the SNP and SSR was carried out to retain highly polymorphic markers and RIL individuals with > 95% genotyping data, as well as markers that exhibited the expected 1:1 segregation ratio.

Linkage Map Construction

Linkage analysis was carried out using the filtered SNP and SSR markers, following Wu et al. (2021). This involved the generation of a draft linkage map using the minimum spanning tree map (MSTMap) (Wu et al., 2008) and then refined by MAPMAKER/EXP 3.0 (Lincoln et al., 1992). The Kosambi map function (Kosambi, 1944) was used to calculate the genetic distances (in cM) between the markers. The map construction was carried out with MapChart v. 2.32 (Voorrips, 2002) using the Kosambi map function, of which the linkage groups were assigned to chromosomes based on the consensus SNP map of pea developed by Tayeh et al. (2015). The sequences of the SNP markers flanking the QTLs associated with partial resistance to FRR caused by *F. graminearum* were used in BlastN (*E*-value \leq E-20) searches of the Pulse Crop Database¹ to determine their possible functions.

Quantitative Trait Loci Analysis

Additive-effect QTL analysis was first carried out using the genotypic and phenotypic data (disease severity, vigor, and plant height) from the RILs inoculated with *F. graminearum* (FG2). This was then repeated for the RILs inoculated with *F. avenaceum* (F4A). The analysis was conducted using means

for the four single greenhouse experiments, LSM, and BLUPs of the total data by Composite Interval Mapping (CIM) using WinQTL Cartographer v2.5 (Wang et al., 2012). The program was set at 1-cM walking speed; forward and backward regression method; window size, 10 cM; five background cofactors; 1,000 permutations, and *p* < 0.05 (Wang et al., 2012). The LOD score threshold was set at 3 for QTL detection. The confidence interval for each QTL was defined by the consensus region bordered by the four environments.

The QTL names were defined according to the QTL detection studies by Coyne et al. (2015, 2019), where the name of the *Fusarium* isolate was indicated, followed by “Ps” = *Pisum sativum*, the first number = the pea linkage group (Tayeh et al., 2015), and the second number = the serial number of the QTL on the linkage group; for example, “Fg-Ps4.1” represents the QTL for disease severity caused by *F. graminearum* located on linkage group IV of the pea genome. The chromosomes and pseudomolecules were named in accordance with Neumann et al. (2002) and Kreplak et al. (2019), respectively. A similar nomenclature was adopted for vigor (e.g., Vig-Ps2.1) and plant height (e.g., Hgt-Ps2.1).

Quantitative trait loci identified in at least two of the four environments were classified as stable. The percentage of variation (R^2) was determined for each QTL. Furthermore, QTL with R^2 > 10%, 5–10%, and <5% were arbitrarily classified as major-, moderate-, or minor-effect QTL, respectively. The origins of favorable alleles for individual traits were assigned to different parents, following Wu et al. (2021). Pairwise epistatic interactions were estimated with IciMapping V.4.1 using the ICIM-EPI method (Meng et al., 2015). The significance threshold for major, moderate, and minor was arbitrarily set at R^2 > 15%, 7.5–15%, and <7.5%, respectively. Epistatic-effect QTL were named with the prefix “E,” followed by the QTL name and a serial number (e.g., E.FG-Ps1, E.Vig-Ps1, and E.Hgt-Ps1).

RESULTS

Preliminary Root Rot Assessment in Parents Against Five *Fusarium* spp.

Between the parental cultivars, “00-2067” developed lower root rot severity than “Reward” in response to each of the five isolates (Supplementary Table 1), confirming that “00-2067” was tolerant, while “Reward” was susceptible. There were significant differences (*p* < 0.001) between the mean root rot values of the tolerant parent “00-2067” and the susceptible parent “Reward”, following inoculation with *F. graminearum* isolate FG2 (Figure 1a) and *F. avenaceum* isolate F4A, while no significant differences were detected following inoculation with the *F. solani*, *F. acuminatum*, and *F. proliferatum* isolates S4C, F037, and F039, respectively (Supplementary Table 1). Therefore, FG2 and F4A were selected to screen the 135 F₈ RIL population for QTL identification associated with resistance to FRR.

ANOVA for Disease Severity, Vigor, and Plant Height

The mean root rot severity, vigor, and plant height of the RIL population inoculated with FG2 and F4A are presented

¹ www.pulsedb.org/

in **Tables 1, 2**. ANOVA indicated that the genotypic effect of disease severity, vigor, and plant height was significant ($p < 0.001$) (**Supplementary Tables 2a,b**). This suggested that a high proportion of genetic variance was transmitted from parental cultivars to the progenies. Heritability values of 92 and 86% for disease severity and vigor were obtained for plants inoculated with FG2 and F4A, respectively, while heritability values for plant height ranged from 79 to 91% (**Supplementary Tables 2a,b**). The $G \times E$ interactions were significant for disease severity, vigor, and plant height for F4A but not for FG2, while differences among the four greenhouse experiments were significant for both FG2 and F4A ($p < 0.001$) (**Supplementary Tables 2a,b**).

Root Rot, Vigor, and Plant Height of Parents and the Recombinant Inbred Line Population Inoculated With FG2

Estimated disease severity values (\pm SE) on the parental cultivar “00-2067” inoculated with FG2 were 1.5 ± 0.7 , 1.3 ± 0.6 , 2 ± 1.2 and 1.5 ± 0.7 for the four greenhouse experiments, 1.6 ± 0.8 for LSM and 1.2 for the BLUPs. This was comparable with the estimated mean of 1.1 ± 0.4 obtained in the preliminary screening of the parents (**Table 1** and **Supplementary Table 1**). On the other hand, the estimated disease severity values (\pm SE) for “Reward” were 3.3 ± 0.5 , 3.3 ± 0.5 , 3.5 ± 0.6 , and 3.0 ± 0.0 for the four greenhouse experiments, 3.3 ± 0.5 for LSM and 4.1 for the BLUPs; these values were also comparable to the estimated mean of 3.3 ± 0.4 obtained in the preliminary screening (**Table 1** and **Supplementary Table 1**). A t -test indicated a significant difference between the parents for disease severity in all four experiments. Frequency distribution (**Figure 2A**) indicated that the disease severity data of the RILs in the four experiments were continuous, but only DSGH3 and DSGHC followed a normal distribution based on the Shapiro–Wilk test (**Table 1**). High correlation coefficients, ranging from 68 to 99%, were found for disease severity among the single experiments, pooled, and BLUPs (**Figure 2A**). The differences in vigor between the parents inoculated with FG2 were significant, except for VGH4. The parental cultivar, “00-2067” had estimated means (\pm SE) of 4.0 ± 0.0 , 4.0 ± 0.0 , 3.0 ± 1.2 and 3.5 ± 0.7 for the four greenhouse experiments and 3.6 ± 0.8 for the pooled data. In the case of “Reward”, the estimated means (\pm SE) were 2.0 ± 0.8 , 2.5 ± 0.6 , 1.5 ± 0.6 and 2.7 ± 0.6 for the four greenhouse experiments and 2.2 ± 0.7 for the pooled data. The BLUPs for the parental cultivars “00-2067” and “Reward” were 4.2 and 1.6, respectively (**Table 1**). The Shapiro–Wilk test indicated that the RIL population vigor data for the four greenhouse experiments did not follow a normal distribution, except for VGH4 (**Figure 2B**). A significant correlation ($0.34 < r < 0.96$, $p < 0.001$) existed among the single experiments, pooled, and BLUPs for vigor (**Figure 2B**). The height of “00-2067” plants inoculated with FG2 was greater than plants of “Reward” for the means in the single environments, LSM, and BLUPs, although the differences were not significant based on a t -test. The estimated means in single conditions, LSM, and BLUP for plant height (\pm SE) of “00-2067” were 234.5 ± 54.6 cm,

157.3 ± 50.6 cm, 155.5 ± 59.5 cm, 159.7 ± 6.7 cm, 176.7 ± 56.6 cm and 158.9 cm, respectively. For “Reward”, the plant heights were 177.5 ± 36.1 cm, 120.7 ± 31.5 cm, 129.5 ± 26 cm, 178.5 ± 34.6 cm, 151.5 ± 36.9 cm, and 100.8 cm, respectively. The frequency distribution of plant height of the RIL population for all six variables was not normal and slightly skewed (**Table 1** and **Figure 2C**). A high correlation ($0.42 < r < 0.95$, $p < 0.001$) was found for plant height among the single experiments, pooled, and BLUPs data (**Figure 2C**).

Collectively, the correlation analysis among traits indicated that root rot caused by FG2 was negatively correlated with vigor and plant height. High correlation coefficients were detected between disease severity and vigor in all conditions ($-0.65 < r < -0.90$, $p < 0.001$), indicating the adverse effect of FG2 on root and aboveground growth. Plant height showed low to moderate correlation with disease severity ($-0.22 < r < -0.35$, $p < 0.05$) and vigor ($0.19 < r < 0.38$, $p < 0.05$).

Root Rot, Vigor, and Plant Height of Parents and the RIL Population Inoculated With F4A

The estimated means (\pm SE) of disease severity for “00-2067” were 1.0 ± 0.0 , 1.0 ± 0.8 , 1.3 ± 0.5 , 1.0 ± 0.0 , 1.1 ± 0.4 , and 1.0, while, for “Reward”, they were 3.3 ± 0.5 , 3.3 ± 0.5 , 3.5 ± 0.6 , 3.0 ± 0.0 , 3.3 ± 0.4 , and 3.0 for DSGH1, DSGH2, DSGH3, DSGH4, LSM of pooled data, and BLUPs, respectively (**Table 2**). These values were comparable to the estimated means (\pm SE) of 1.8 ± 0.5 and 2.8 ± 0.2 for disease severity obtained in the preliminary screening of the parents (**Table 2** and **Supplementary Table 1**). t -tests indicated significant differences between estimated means of the parental cultivars “00-2067” and “Reward” inoculated with F4A. The Shapiro–Wilk test indicated that only the root rot data of the RIL population for DSGH4 and DSGH Pooled followed a normal distribution (**Table 2**), although the data for the four greenhouse experiments were continuous (**Figure 3A**). The correlation coefficient between the experiments ranged from 0.44 to 0.93 ($p < 0.001$) (**Figure 3A**). Based on the t -tests, the parental cultivar “00-2067” inoculated with F4A had significantly greater vigor than “Reward.” The estimated vigor values (\pm SE) for “00-2067” were 4.0 ± 0.0 , 3.7 ± 0.5 , 3.5 ± 0.6 , and 4 ± 0 for the four individual greenhouse experiments, 3.8 ± 0.5 for LSM for the pooled data and 4.0 for BLUPs of the pooled greenhouse experiments (**Table 2**). The estimated vigor values (\pm SE) for “Reward” were 1.7 ± 0.5 , 2.0 ± 1.4 , 1.2 ± 1.5 , and 2.5 ± 0.6 for the individual greenhouse experiments, 1.9 ± 1.1 for LSM, and 1.9 for the BLUPs of the pooled greenhouse experiments. All vigor variables for the RIL population were continuous with slight left skewness ($-0.4 \sim -0.9$) (**Figure 3B**). Additionally, the data did not follow a normal distribution based on the Shapiro–Wilk test (**Table 2**). The correlation coefficient between the experiments ranged from 0.54 to 0.98 ($p < 0.001$) (**Figure 3B**).

In contrast to vigor, the difference in plant height of the parental cultivars inoculated with F4A was not significant based on the t -test. The estimated plant height for “00-2067” for the individual experiments, LSM, and BLUP was 210.8 ± 128.2 cm, 174.5 ± 104.8 cm, 159.5 ± 13.5 cm, 210.0 ± 53.1 cm,

TABLE 1 | A statistical summary of phenotypic data for the parental pea cultivars, “00-2067” and “Reward”, and an RIL population inoculated with *Fusarium graminearum* isolate FG2, in four greenhouse experiments, as well as the pooled and the best linear unbiased predictors (BLUPs).

| Trait | Abbrev/Experiment | Parental cultivar | | | RIL population | | | |
|-------------------|-------------------|------------------------|-----------------------|-----------------------------|-------------------|----------|----------|---------------------------|
| | | ‘00-2067’ ^a | ‘Reward’ ^a | <i>T</i> -test (<i>P</i>) | RILs ^a | Skewness | Kurtosis | Shapiro-test (<i>P</i>) |
| Root rot severity | DSGH1 | 1.5 ± 0.7 | 3.3 ± 0.5 | 1.1E-02 | 2.0 ± 0.7 | −0.1 | −0.7 | 1.6E-03 |
| Root rot severity | DSGH2 | 1.3 ± 0.6 | 3.3 ± 0.5 | 2.6E-03 | 1.9 ± 0.7 | −0.2 | −0.6 | 1.5E-02 |
| Root rot severity | DSGH3 | 2.0 ± 1.2 | 3.5 ± 0.6 | 3.0E-02 | 2.2 ± 0.7 | 0.0 | −0.2 | 6.0E-02 |
| Root rot severity | DSGH4 | 1.5 ± 0.7 | 3.0 ± 0.0 | 1.3E-02 | 2.1 ± 0.7 | 0.1 | −0.3 | 3.7E-02 |
| Root rot severity | DSGHPooled | 1.6 ± 0.8 | 3.3 ± 0.5 | 4.6E-07 | 2.0 ± 0.6 | −0.1 | −0.7 | 6.2E-02 |
| Root rot severity | DSGHBLUPS | 1.2 | 4.1 | – | 2.0 ± 1.2 | −0.1 | −0.8 | 4.6E-02 |
| Vigor | VGH1 | 4.0 ± 0.0 | 2.0 ± 0.8 | 1.5E-02 | 3.0 ± 0.8 | −0.3 | −0.7 | 0.0E + 00 |
| Vigor | VGH2 | 4.0 ± 0.0 | 2.5 ± 0.6 | 3.5E-03 | 3.0 ± 0.7 | −0.3 | −0.2 | 9.5E-06 |
| Vigor | VGH3 | 3.0 ± 1.2 | 1.5 ± 0.6 | 3.0E-02 | 2.7 ± 0.7 | −0.2 | −0.2 | 5.1E-06 |
| Vigor | VGH4 | 3.5 ± 0.7 | 2.7 ± 0.6 | 1.2E-01 | 2.8 ± 0.8 | −0.4 | 0.6 | 1.1E-03 |
| Vigor | VGHPooled | 3.6 ± 0.8 | 2.2 ± 0.7 | 6.0E-05 | 2.9 ± 0.5 | 0.0 | −0.5 | 1.1E-01 |
| Vigor | VGHBLUPS | 4.2 | 1.6 | – | 2.9 ± 1.1 | −0.1 | −0.4 | 5.1E-02 |
| Plant height | HGH1 | 234.5 ± 54.6 | 177.5 ± 36.1 | 1.3E-01 | 217.6 ± 96.3 | 1.0 | 0.7 | 0.0E + 00 |
| Plant height | HGH2 | 157.3 ± 50.6 | 120.7 ± 31.5 | 1.6E-01 | 231.5 ± 87.4 | 0.7 | 0.2 | 5.6E-05 |
| Plant height | HGH3 | 155.5 ± 59.5 | 129.5 ± 26.0 | 2.3E-01 | 154.5 ± 84.4 | 1.0 | 1.1 | 8.0E-06 |
| Plant height | HGH4 | 159.7 ± 6.7 | 178.5 ± 34.6 | 2.0E-01 | 184.6 ± 83.9 | 0.6 | 0.5 | 5.5E-02 |
| Plant height | HGHPooled | 176.7 ± 56.6 | 151.5 ± 36.9 | 5.1E-02 | 197.5 ± 68.8 | 1.1 | 1.2 | 0.0E + 00 |
| Plant height | HGHBLUPS | 158.9 | 100.8 | – | 197.3 ± 135.5 | 1.0 | 0.6 | 0.0E + 00 |

^aThe estimated means of parental cultivars, “Reward” and “00-2067,” as well as RILs, are shown along with plus/minus stand error (SE).

TABLE 2 | A statistical summary of phenotypic data for the parental pea cultivars, “00-2067” and “Reward”, and an RIL population inoculated with *Fusarium avenaceum* isolate F4A, in four greenhouse experiments, as well as the pooled and the best linear unbiased predictors (BLUPs) of the greenhouse experiments.

| Trait | Abbrev/Experiment | Parental cultivar | | | RIL population | | | |
|-------------------|-------------------|------------------------|-----------------------|-----------------------------|-------------------|----------|----------|---------------------------|
| | | ‘00-2067’ ^a | ‘Reward’ ^a | <i>T</i> -test (<i>P</i>) | RILs ^a | Skewness | Kurtosis | Shapiro-test (<i>P</i>) |
| Root rot severity | DSGH1 | 1.0 ± 0.0 | 3.3 ± 0.5 | 5.30E-05 | 2.2 ± 0.9 | 0.2 | −0.9 | 7.77E-05 |
| Root rot severity | DSGH2 | 1.0 ± 0.8 | 3.3 ± 0.5 | 1.66E-03 | 2.3 ± 0.9 | 0.3 | −0.6 | 1.35E-04 |
| Root rot severity | DSGH3 | 1.3 ± 0.5 | 3.5 ± 0.6 | 5.30E-04 | 2.5 ± 0.9 | 0.0 | −1.0 | 2.86E-06 |
| Root rot severity | DSGH4 | 1.0 ± 0 | 3.0 ± 0.0 | 1.36E-03 | 2.4 ± 0.9 | 0.0 | −0.7 | 1.38E-02 |
| Root rot severity | DSGHPooled | 1.1 ± 0.4 | 3.3 ± 0.4 | 2.63E-13 | 2.4 ± 0.7 | 0.0 | −0.6 | 1.20E-01 |
| Root rot severity | DSGHBLUPS | 1.0 | 3.0 | – | 2.3 ± 0.9 | 0.1 | −1.0 | 8.07E-05 |
| Vigor | VGH1 | 4.0 ± 0.0 | 1.7 ± 0.5 | 5.26E-05 | 2.6 ± 1.1 | −0.5 | −0.6 | 5.96E-08 |
| Vigor | VGH2 | 3.7 ± 0.5 | 2.0 ± 1.4 | 2.92E-02 | 2.6 ± 1.1 | −0.9 | −0.1 | 0.00E + 00 |
| Vigor | VGH3 | 3.5 ± 0.6 | 1.2 ± 1.5 | 1.56E-02 | 2.4 ± 1.2 | −0.5 | −0.7 | 0.00E + 00 |
| Vigor | VGH4 | 4.0 ± 0.0 | 2.5 ± 0.6 | 1.01E-03 | 2.5 ± 1.1 | −0.4 | −1.0 | 1.79E-07 |
| Vigor | VGHPooled | 3.8 ± 0.5 | 1.9 ± 1.1 | 1.06E-07 | 2.5 ± 0.9 | −0.5 | −0.6 | 9.89E-06 |
| Vigor | VGHBLUPS | 4.0 | 1.9 | – | 2.6 ± 1 | −0.6 | −0.3 | 5.96E-08 |
| Plant height | HGH1 | 210.8 ± 128.2 | 118.3 ± 100.2 | 1.49E-01 | 196.8 ± 87.9 | 0.7 | 0.6 | 1.35E-03 |
| Plant height | HGH2 | 174.5 ± 104.8 | 193.5 ± 104.5 | 4.03E-01 | 194.5 ± 88.5 | 0.6 | 1.3 | 9.74E-02 |
| Plant height | HGH3 | 159.5 ± 13.5 | 125.0 ± 98.9 | 2.58E-01 | 171.8 ± 80 | 0.4 | 0.3 | 3.76E-01 |
| Plant height | HGH4 | 210.0 ± 53.1 | 194.0 ± 38.4 | 3.00E-01 | 180.8 ± 80.1 | 0.7 | 0.7 | 8.17E-03 |
| Plant height | HGHPooled | 188.7 ± 82.8 | 157.7 ± 88.5 | 1.54E-01 | 185.2 ± 64 | 0.6 | 0.5 | 9.79E-02 |
| Plant height | HGHBLUPS | 208.0 | 133.1 | – | 186.6 ± 122.6 | 0.6 | 1.0 | 5.72E-02 |

^aThe estimated means of parental cultivars, “Reward” and “00-2067,” as well as RILs, are shown along with plus/minus stand error (SE).

188.7 ± 82.8 cm, and 208.0 cm, respectively. The estimated plant height for “Reward” was 118.3 ± 100.2 cm, 193.5 ± 104.5 cm, 125. ± 98.9 cm, and 194.0 ± 38.4 cm for the individual experiments, 157.7 ± 88.5 cm for LSM and 133.1 cm for the BLUP. The frequency distribution for the RIL population

was continuous and slightly skewed to the right. In addition, HGH2, HGH3, HGH Pooled, and HGH BLUPs followed a normal distribution (Table 2 and Figure 3C). Plant height variables were also significantly correlated ($0.28 < r < 0.97$, $p < 0.01$) (Figure 3C).

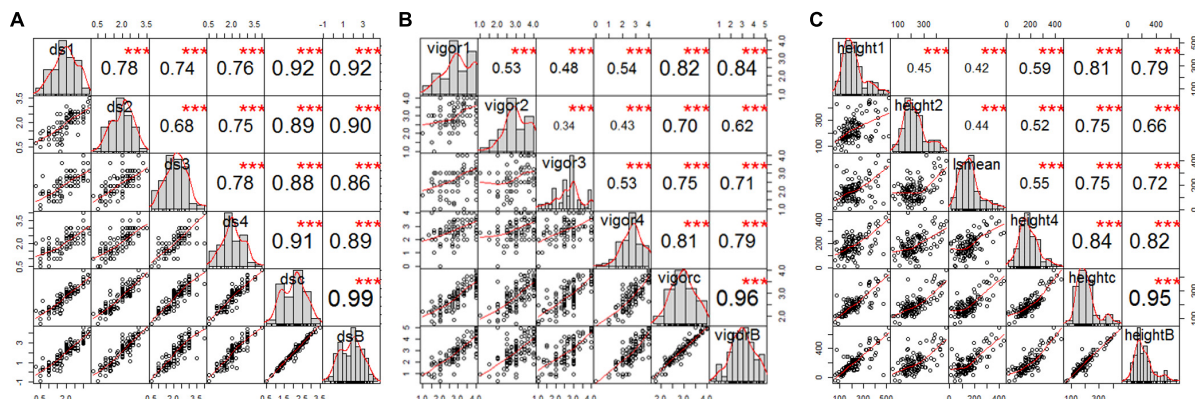


FIGURE 2 | Correlation analysis of estimated mean of four single greenhouse experiments, BLUPs, and combined total data for (A) root rot severity, (B) vigor, and (C) height of pea inoculated with FG2, illustrating the significant correlation among all variables for each trait. The bar graphs indicate the frequency distributions across the diagonal. The correlation coefficients with a significance level (* indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$) and scatter plots between pairs are shown above and below the diagonal, respectively.

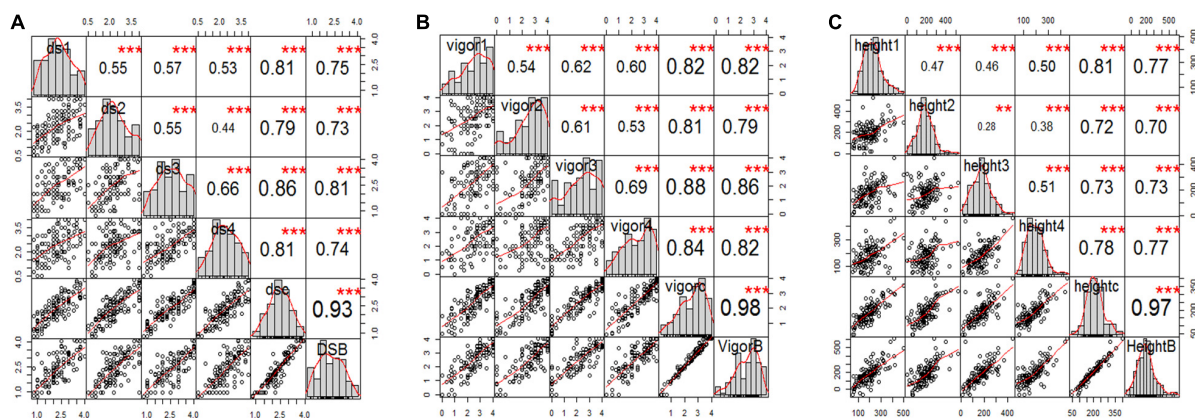


FIGURE 3 | Correlation analysis among three pea root rot disease-related traits of (A) root rot severity, (B) vigor, and (C) plant height for all six variables, including the estimated mean of four single greenhouse studies, combined total data, and BLUPs (e.g., panels ds1, ds2, ds3, ds4, dsc, and dsB for disease severity) inoculated with F4A. The bar graphs indicate the frequency distributions across the diagonal. The correlation coefficients with a significance level (* indicates $p < 0.05$; ** indicates $p < 0.01$; *** indicates $p < 0.001$) and scatter plots between pairs are shown above and below the diagonal, respectively.

The correlation among the traits for plants inoculated with F4A was similar to that of plants inoculated with FG2. Disease severity was highly correlated with vigor ($-0.88 < r < -0.95$, $p < 0.001$) and with plant height ($-0.48 < r < -0.63$, $p < 0.001$). Plant height was positively correlated with vigor ($0.57 < r < 0.62$, $p < 0.001$).

Genetic Map Construction and Quantitative Trait Loci Analysis

Linkage grouping, the distribution of markers, map length, and marker density of 2999 (2978 SNP + 21 SSR) retained markers were as described by Wu et al. (2021). The marker distribution in this study was compared with the seven chromosomes of pea as determined by Neumann et al. (2002), linkage groups as determined by Tayeh et al. (2015), and pseudomolecules of pea (Kreplak et al., 2019). The genetic map spanned 1704.1 cM and contained an average marker density of 1.8 markers/cM

(Wu et al., 2021). The QTL analysis was conducted with 1,422 unique markers, which represented 10.5% of the markers used for genotyping (Wu et al., 2021).

Additive-Effect Quantitative Trait Loci Analysis

No significant QTL ($LOD > 3.0$) for disease severity, vigor, and plant height were detected for the RILs inoculated with *F. avenaceum* isolate F4A. As such, no QTL likelihood profiles are shown. In the case of RILs inoculated with *F. graminearum* isolate FG2, a total of 11 QTL were detected for the three parameters and six variables (i.e., GH1, GH2, GH3, GH4, LSM, and BLUPs) by the CIM using Win QTL Cartographer v2.5 (Wang et al., 2012; Table 3). Five of the 11 QTL were identified for disease severity, whereas three QTL each were detected for vigor and plant height. The QTL had LOD scores ranging from 3.0 to 14.4 and the percentage of phenotypic variation (R^2) values ranging

TABLE 3 | A summary of the QTL associated with *Fusarium* root rot severity, vigor, and plant height in 128 F₈-derived recombinant inbred pea lines from the cross between the cultivars “Reward” × “00-2067” under greenhouse (GH) conditions.

| Identified QTL | Trait | Environment | LG Analysis (Present study) | Chrom ^a /LG ^b | Peak (cM) | Confidence interval(cM) | Left Marker | Right marker | LOD | Additive | R ² (%) |
|------------------|-------------------|-------------|-----------------------------|-------------------------------------|-----------|-------------------------|------------------------|------------------------|------|----------|--------------------|
| <i>Fg-Ps3.1</i> | Root rot severity | GH Expt 1 | 4 | Chrom5/LGIII | 311.9 | 307.9–316.5 | AA5 | PsCam036163_21311_1095 | 3.9 | −0.2409 | 9.88 |
| <i>Fg-Ps3.2</i> | Root rot severity | GH Expt 4 | 4 | Chrom5/LGIII | 338.2 | 334.9–341.4 | PsCam036163_21311_1095 | PsCam042783_26826_1395 | 3.5 | −0.2153 | 9.62 |
| <i>Fg-Ps4.1</i> | Root rot severity | GH Expt 1 | 5 | Chrom4/LGIV | 71.7 | 63.7–74.4 | PsCam050913_33466_1250 | PsCam048871_31524_450 | 3.0 | −0.2121 | 9.10 |
| | Root rot severity | GH Expt 2 | 5 | Chrom4/LGIV | 61.3 | 59.3–69.2 | PsCam001381_1152_437 | PsCam042375_26443_3427 | 3.8 | −0.2229 | 10.57 |
| <i>Fg-Ps4.2</i> | Root rot severity | GH Expt 3 | 5 | Chrom4/LGIV | 80.2 | 74.0–85.2 | AA239 | PsCam057281_37909_2940 | 5.9 | −0.2344 | 11.26 |
| | Root rot severity | GH Expt 4 | 5 | Chrom4/LGIV | 80.2 | 75.4–85.2 | AA239 | PsCam057281_37909_2940 | 4.1 | −0.2526 | 13.17 |
| | Root rot severity | Pooled | 5 | Chrom4/LGIV | 79.2 | 75.4–85.2 | AA239 | PsCam057281_37909_2940 | 5.1 | −0.2492 | 15.44 |
| | Root rot severity | BLUPs | 5 | Chrom4/LGIV | 79.2 | 75.4–85.2 | AA239 | PsCam057281_37909_2940 | 3.7 | −0.3838 | 10.02 |
| <i>Fg-Ps5.1</i> | Root rot severity | GH Expt 1 | 7 | Chrom3/LGV | 5.2 | 0.9–9.2 | PsCam059449_39630_321 | PsCam011153_7569_125 | 5.5 | 0.3036 | 14.22 |
| <i>Vig-Ps3.1</i> | Vigor | GH Expt 4 | 4 | Chrom5/LGIII | 68.9 | 67.1–70.5 | PsCam013763_9362_423 | AD270 | 4.9 | −0.1423 | 4.05 |
| <i>Vig-Ps3.2</i> | Vigor | GH Expt 2 | 4 | Chrom5/LGIII | 312.0 | 307.8–316.8 | AA5 | PsCam036163_21311_1095 | 3.0 | 0.1910 | 9.53 |
| | Vigor | GH Expt 3 | 4 | Chrom5/LGIII | 316.1 | 310.4–320.4 | AA5 | PsCam036163_21311_1095 | 3.3 | 0.2582 | 11.22 |
| | Vigor | Pooled | 4 | Chrom5/LGIII | 312.6 | 310.4–316.5 | AA5 | PsCam036163_21311_1095 | 4.6 | 0.1938 | 12.13 |
| | Vigor | BLUPs | 4 | Chrom5/LGIII | 312.6 | 307.5–316.5 | AA5 | PsCam036163_21311_1095 | 4.2 | 0.3736 | 11.92 |
| | Vigor | GH Expt 1 | 5 | Chrom4/LGIV | 68.0 | 63.5–69.9 | PsCam050913_33466_1250 | PsCam042375_26443_3427 | 4.4 | 0.2728 | 13.50 |
| <i>Vig-Ps4.1</i> | Vigor | GH Expt 2 | 5 | Chrom4/LGIV | 60.3 | 58.0–70.7 | PsCam000712_620_237 | PsCam042375_26443_3427 | 3.2 | 0.2060 | 10.42 |
| | Vigor | GH Expt 4 | 5 | Chrom4/LGIV | 71.5 | 70.5–73.2 | PsCam042375_26443_3427 | AA239 | 4.5 | 0.2437 | 9.19 |
| | Vigor | Pooled | 5 | Chrom4/LGIV | 61.3 | 58.8–63.5 | PsCam000712_620_237 | PsCam057555_38139_296 | 4.4 | 0.1868 | 11.59 |
| | Vigor | BLUPs | 5 | Chrom4/LGIV | 61.3 | 59.3–63.5 | PsCam001381_1152_437 | PsCam057555_38139_296 | 3.8 | 0.3474 | 10.51 |
| | Vigor | GH Expt 1 | 4 | Chrom5/LGIII | 288.6 | 288.3–291.7 | PsCam020937_11699_2576 | AA5 | 14.4 | −62.31 | 36.35 |
| <i>Hgt-Ps3.1</i> | Height | GH Expt 2 | 4 | Chrom5/LGIII | 287.6 | 286.8–293.7 | PsCam020937_11699_2576 | AA5 | 4.7 | −33.90 | 12.90 |
| | Height | GH Expt 4 | 4 | Chrom5/LGIII | 287.6 | 286.8–295.2 | PsCam020937_11699_2576 | AA5 | 3.3 | −27.27 | 9.94 |
| | Height | Pooled | 4 | Chrom5/LGIII | 287.6 | 286.8–292.4 | PsCam020937_11699_2576 | AA5 | 6.2 | −33.24 | 20.96 |
| | Height | BLUPs | 4 | Chrom5/LGIII | 287.6 | 286.8–291.4 | PsCam020937_11699_2576 | AA5 | 9.4 | −71.63 | 23.97 |
| | Height | GH Expt 1 | 9 | Chrom7/LGVII | 92.2 | 85.3–102.1 | PsCam039854_24711_656 | PsCam046792_30096_853 | 10.1 | 46.60 | 20.04 |
| <i>Hgt-Ps7.1</i> | Height | Pooled | 9 | Chrom7/LGVII | 92.2 | 84.5–115.3 | PsCam056683_37453_248 | PsCam021891_12310_347 | 4.9 | 28.68 | 13.54 |
| | Height | BLUPs | 9 | Chrom7/LGVII | 92.2 | 81.2–102.5 | PsCam035831_20992_561 | PsCam042171_26273_1937 | 4.5 | 52.40 | 7.04 |
| | Height | GH Expt 2 | 9 | Chrom7/LGVII | 154.3 | 143.8–167.5 | PsCam002756_2184_427 | PsCam045262_28962_162 | 5.1 | 34.63 | 13.63 |
| <i>Hgt-Ps7.2</i> | Height | GH Expt 4 | 9 | Chrom7/LGVII | 144.3 | 142.3–151.9 | PsCam002756_2184_427 | PsCam011213_7616_1104 | 4.4 | 32.77 | 13.85 |
| | Height | Pooled | 9 | Chrom7/LGVII | 157.7 | 148.8–168.0 | AB91 | PsCam045262_28962_162 | 4.4 | 27.87 | 14.06 |

^aPea chromosomes named according to Neumann et al. (2002) and ^bPea linkage groups named according to Tayeh et al. (2015).

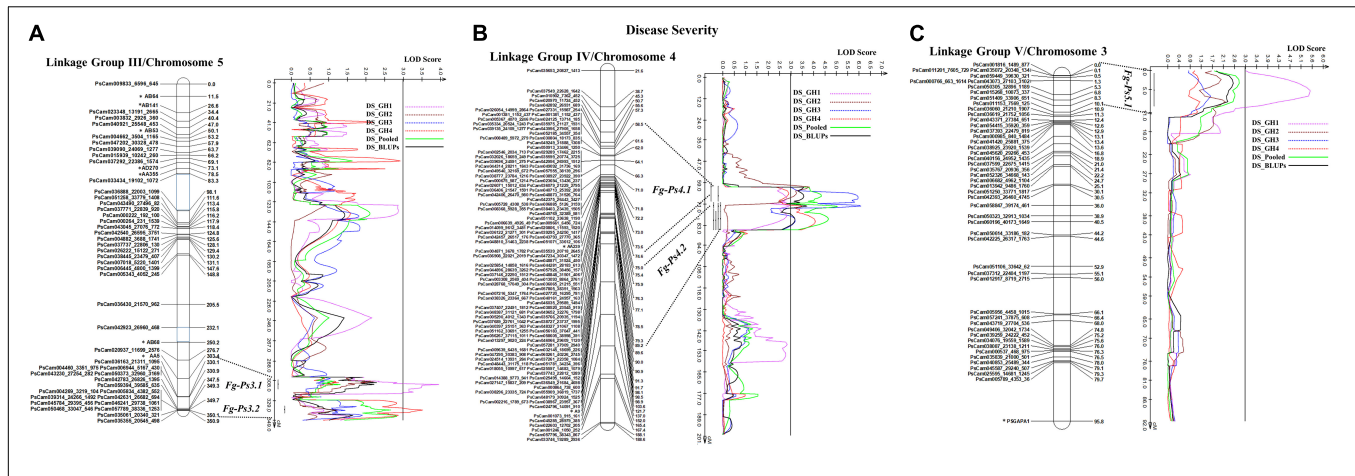


FIGURE 4 | Identified QTL and the linkage map of pea LG III (Chrom 5), IV (Chrom 4), and V (Chrom 3) associated with partial resistance to *Fusarium graminearum* in an F_8 RIL derived from “Reward” \times “002067.” The LOD scores are indicated on the x-axis, while the genetic distances (in cM) are indicated on the y-axis. **(A)** Two minor-effect QTL, *Fg-Ps3.1* and *Fg-Ps3.2*, on LG III (Chrom5) were detected in greenhouse Experiments 1 and 4, respectively. **(B)** Two stable, moderate-effect QTL, *Fg-Ps4.1* and *Fg-Ps4.2*, were located on LG IV (Chrom4) and identified in greenhouse Experiments 1 and 2 and 3 and 4, respectively. **(C)** Another moderate-effect QTL, *Fg-Ps5.1*, on LG V (Chrom5) was detected only in greenhouse Experiment 1.

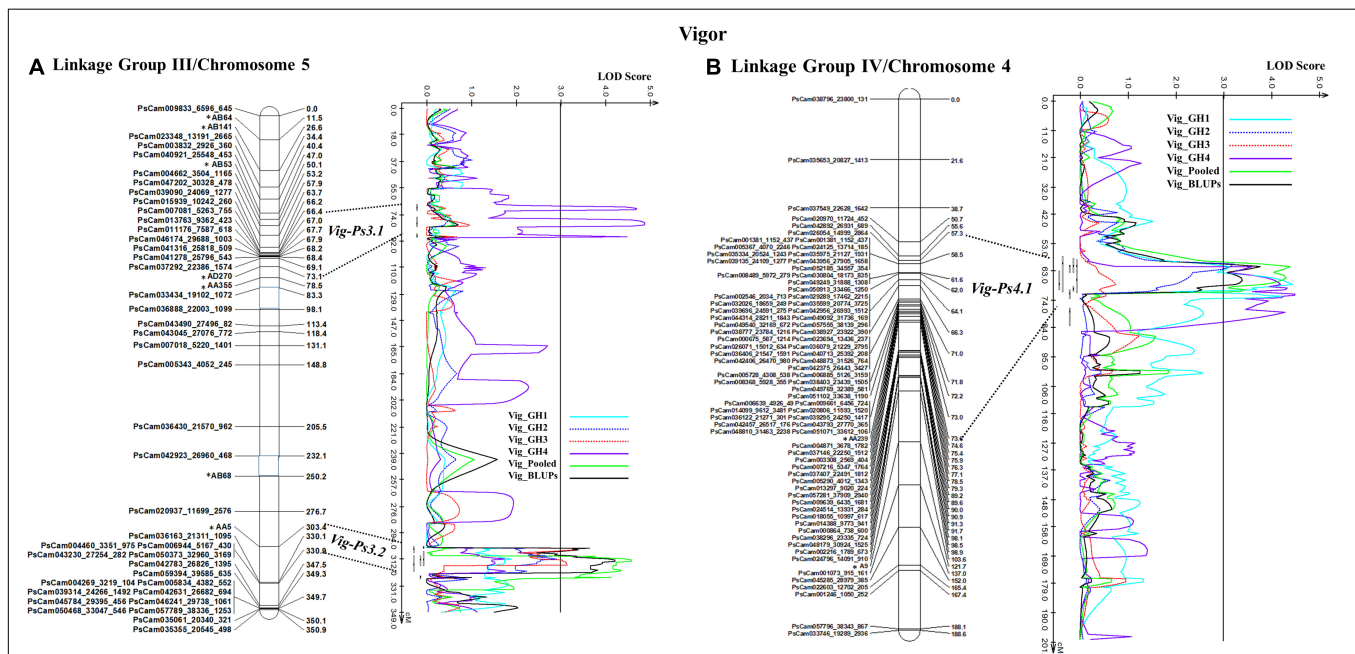
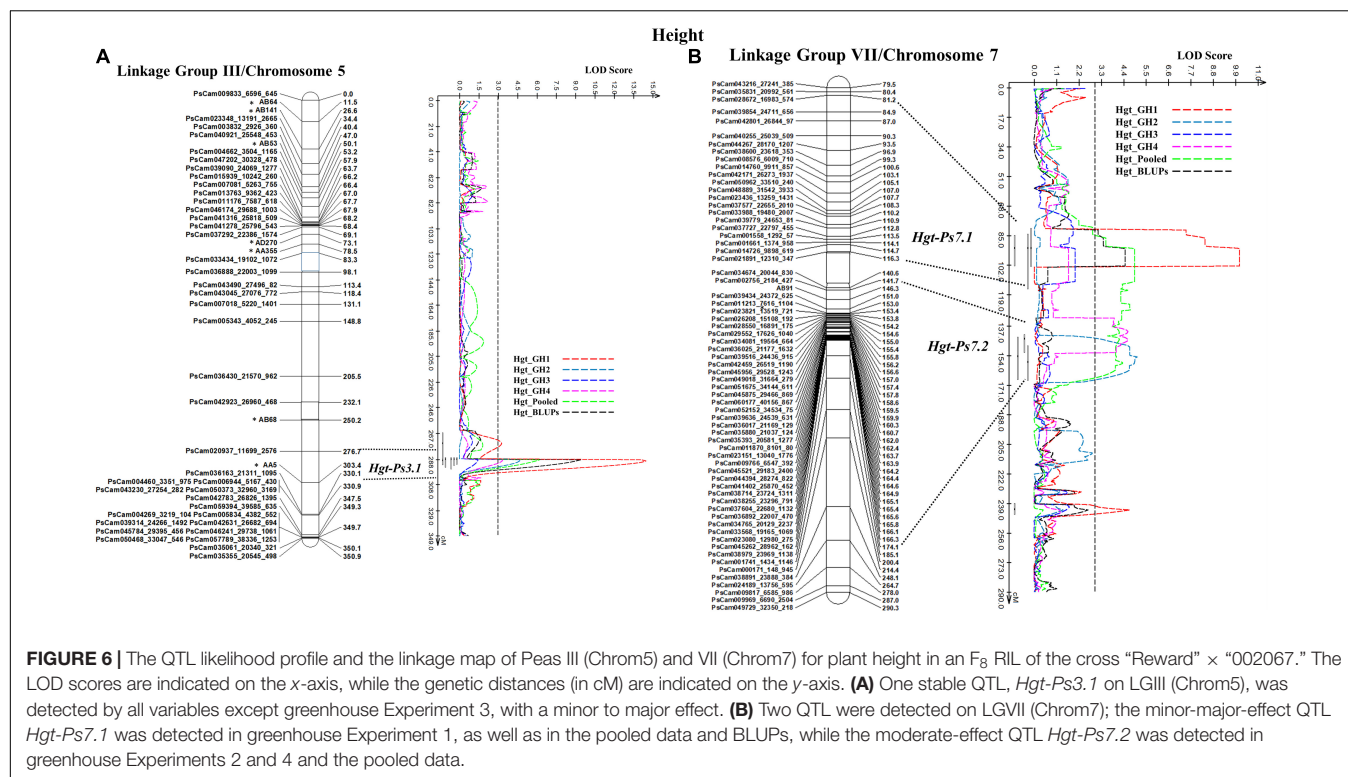


FIGURE 5 | The QTL likelihood profile and the linkage map of pea LG III (Chrom5) and IV (Chrom 4) for vigor in an F_8 RIL of the cross “Reward” \times “002067.” The LOD scores are indicated on the x-axis, while the genetic distances (in cM) are indicated on the y-axis. **(A)** Minor-effect QTL *Vig-Ps3.1* on LG III (Chrom5) was detected for vigor only in greenhouse Experiment 4. Another QTL *Vig-Ps3.2* was identified multiple times in greenhouse Experiments 2 and 3, as well as in the BLUPs and pooled data. **(B)** One minor-moderate-effect QTL, *Vig-Ps4.1*, was identified on LG IV (Chrom 4) greenhouse Experiments 1, 2, and 5, as well as the pooled data and BLUPs.

from 4.05 to 36.35% (Table 3). Based on the R^2 values, two, six, and three of the QTL were considered major, moderate, or minor effect, respectively. Six of the 11 QTL were identified in two or more environments and hence could be considered stable, while the remaining five QTL were detected in single experiments and hence could be considered unstable.

The most stable QTL for partial resistance to *F. graminearum* isolate FG2, *Fg-Ps4.1*, and *Fg-Ps4.2* were located in the

middle of Chrom4/LGIV at positions 59.3–74.4 cM and 74.0–85.2 cM, respectively (Figure 4B). The 15.1-cM and 11.2-cM genomic regions delimiting these two QTL were flanked by the SNP markers PsCam048871_31524_450 and PsCam001381_1152_437 and the SSR marker AA239 and SNP marker PsCam057281_37909_2940, respectively (Figure 4B). Both *Fg-Ps4.1* and *Fg-Ps4.2* exhibited a moderate effect, with the percentage variance ranging from 9.1 to 15.4%



(Table 3). Two other moderate-effect but unstable QTLs, *Fg-Ps3.1* (located on the bottom segment (307.9–316.5 cM) of Chrom5/LGIII and with flanking markers of AA5 and PsCam036163_21311_1095) and *Fg-Ps3.2* (located distal to *Fg-Ps3.1* and with flanking markers PsCam036163_21311_1095 and PsCam042783_26826_1395) explained 9.62–9.88% of the total variance (Figure 4A). Another unstable QTL, *Fg-Ps5.1* [detected on the top part (0.9–9.2 cM) of Chrom3/LGV and flanked by the SNP markers PsCam059449_39630_321 and PsCam011153_7569_125] explained 14.2% of the total variance in greenhouse Experiment 1 (Figure 4C). Four of the QTL for disease severity (with the exception of *Fg-Ps5.1*) had a negative additive effect, indicating that genomic regions for resistance in *Fg-Ps4.1*, *Fg-Ps4.2*, *Fg-Ps3.1*, and *Fg-Ps3.2* originated from “00-2067,” while *Fg-Ps5.1* derived its resistance from “Reward” (Table 3).

The stability of the QTL for vigor was in the order *Vig-Ps4.1* on Chrom4/LGIV (GH1, GH2, and GH4, $R^2 = 9.19$ to 13.5%) $>$ *Vig-Ps3.2* (GH2 and GH3, $R^2 = 9.53\%$ to 12.13%) $>$ *Vig-Ps3.1* (GH4, $R^2 = 4.05\%$) both on Chrom5/LGIII (Table 3). The QTL *Vig-Ps4.1* was located on Chrom4/LGIV from 58.0 cM to 73.2 cM between the SNP marker PsCam000712_620_237 and the SSR marker AA239 (Figure 5B). *Vig-Ps3.2*, which was located 307.8–316.5 cM on the bottom of Chrom5/LGIII, was flanked by the SSR marker AA5 and the SNP marker PsCam036163_21311_1095; *Vig-Ps3.1*, which was located on the top segment (67.1–70.5 cM) of the same chromosome or linkage group, was flanked by the SNP marker PsCam013763_9362_423 and the SSR marker AD270 (Figure 5). The two stable QTL, *Vig-Ps4.1* and *Vig-Ps3.2*, had a positive additive effect, indicating that the alleles for vigor originated from

“00-2067.” In contrast, *Vig-Ps3.1* has a negative additive effect, indicating that the alleles originated from “Reward” (Table 3).

In the case of plant height, the most stable QTL, *Hgt-Ps3.1*, was detected in three of the four experiments (GH1, GH2, and GH4; $R^2 = 9.94$ –36.35%). This QTL was located on the bottom segment of Chrom5/LGIII (Figure 6) and was flanked by the SNP marker PsCam020937_11699_2576 and the SSR marker AA5 (Figure 6A). The second most stable QTL, *Hgt-Ps7.2*, was detected across two (GH2 and GH4) of four greenhouse experiments ($R^2 = 7.04$ –20.04%). These QTL were located 142.3–168.0 cM on Chrom7/LGVII and were flanked by the SNP markers PsCam002756_2184_427 and PsCam045262_28962_162 (Table 3 and Figure 6B). *Hgt-Ps7.1*, which was flanked by the SNP markers PsCam035831_20992_561 and PsCam021891_12310_347 (81.2–115.3 cM) (Figure 6B), was detected in only one environment (GH1) on the same chromosome ($R^2 = 13.63$ –14.06). The additive effect was negative for *Hgt-Ps3.1*, but positive for *Hgt-Ps7.1* and *Hgt-Ps7.2* (Table 3). This suggested that the QTL for height on Chrom5/LGIII was derived from “Reward”, while the QTL on Chrom7/LGVII originated from “00-2067.”

Epistatic Quantitative Trait Loci Analyses

Two hundred eight putative digenic epistatic pairs were identified using all variables for disease severity, vigor, and plant height. These comprised 65 (12–24) for disease severity, 57 (10–21) for vigor, and 86 (15–28) for plant height. The 208 putative digenic interactions consisted of one major epistatic effect ($PVE \geq 15\%$), 13 moderate epistatic effects ($7.5\% \leq PVE \leq 15\%$), and 194 minor epistatic effects ($PVE \leq 7.5\%$). BLUPs for disease severity,

Table of contents

- 07 **Editorial: Advances in Breeding for Quantitative Disease Resistance**
Valerio Hoyos-Villegas, Jianjun Chen, Anna Maria Mastrangelo and Harsh Raman
- 12 **Characterization of the Genetic Architecture for Fusarium Head Blight Resistance in Durum Wheat: The Complex Association of Resistance, Flowering Time, and Height Genes**
Yuefeng Ruan, Wentao Zhang, Ron E. Knox, Samia Berraies, Heather L. Campbell, Raja Ragupathy, Kerry Boyle, Brittany Polley, Maria Antonia Henriquez, Andrew Burt, Santosh Kumar, Richard D. Cuthbert, Pierre R. Fobert, Hermann Buerstmayr and Ron M. DePauw
- 29 **Genomic Breeding for Diameter Growth and Tolerance to *Leptocybe* Gall Wasp and *Botryosphaeria/Teratosphaeria* Fungal Disease Complex in *Eucalyptus grandis***
Makobatjatji M. Mphahlele, Fikret Isik, Gary R. Hodge and Alexander A. Myburg
- 44 **Dissecting Quantitative Trait Loci for Spot Blotch Resistance in South Asia Using Two Wheat Recombinant Inbred Line Populations**
Chandan Roy, Navin C. Gahtyari, Xinyao He, Vinod K. Mishra, Ramesh Chand, Arun K. Joshi and Pawan K. Singh
- 53 **Genome-Wide Association Studies Reveal All-Stage Rust Resistance Loci in Elite Durum Wheat Genotypes**
Meriem Aoun, Matthew N. Rouse, James A. Kolmer, Ajay Kumar and Elias M. Elias
- 73 **Genetic Improvement for Resistance to Black Sigatoka in Bananas: A Systematic Review**
Julianna M. S. Soares, Anelita J. Rocha, Fernanda S. Nascimento, Adriadna S. Santos, Robert N. G. Miller, Cláudia F. Ferreira, Fernando Haddad, Vanusia B. O. Amorim and Edson P. Amorim
- 88 **Introgression of Maize Lethal Necrosis Resistance Quantitative Trait Loci Into Susceptible Maize Populations and Validation of the Resistance Under Field Conditions in Naivasha, Kenya**
Luka A. O. Awata, Beatrice E. Ifie, Eric Danquah, MacDonald Bright Jumbo, L. Mahabaleswara Suresh, Manje Gowda, Philip W. Marchelo-Dragga, Michael Scott Olsen, Oluwaseyi Shorinola, Nasser Kouadio Yao, Prasanna M. Boddupalli and Pangirayi B. Tongoona
- 102 **Novel Genomic Regions of *Fusarium* Wilt Resistance in Bottle Gourd [*Lagenaria siceraria* (Mol.) Standl.] Discovered in Genome-Wide Association Study**
Yanwei Li, Ying Wang, Xinyi Wu, Jian Wang, Xiaohua Wu, Baogen Wang, Zhongfu Lu and Guojing Li

vigor, and plant height detected 20, 18, and 21 putative digenic interactions, respectively. Within the 59 digenic interactions among BLUPs of all traits, epistatic analysis identified one major QTL pair, three moderate QTL pairs, and 55 minor QTL pairs (Table 4). In contrast, LSM of the pooled data detected 23 digenic interactions for disease severity, 14 for vigor, and 19 for plant height. The total 56 pairs included seven moderate epistatic-effect QTLs and 49 minor-effect QTLs.

Twenty-five digenic epistatic interactions with major and moderate effects were identified by 33 flanking markers, of which 10 epistatic-effect QTL with 14 flanking markers were linked to three additive-effect QTL (*Fg-Ps3.1*, *Fg-Ps3.2*, and *Vig-Ps3.1*). The remaining 15 epistatic QTL were not related to any of the additive-effect QTL (Table 4). Eight of the 10 epistatic-effect QTL were linked to *Fg-Ps3.2*, including the most significant QTL pairs, *E.Hgt-Ps1* ($R^2 = 31.2\%$), followed by *E.Hgt-Ps7* ($R^2 = 19.1\%$) and *E.Hgt-Ps4* ($R^2 = 13.5\%$). The fourth was *E.Hgt-Ps3* ($R^2 = 13.5\%$), which was linked to *Fg-Ps3.2* and *Vig-Ps3.1*. Only *E.Fg-Ps7* and *E.Vig-Ps7* were linked to *Fg-Ps3.1*, showing moderate epistatic effect ($R^2 = 9.5\%$ and $R^2 = 12.6\%$, respectively).

Candidate Genes

The QTL associated with partial resistance to *F. graminearum* on Chrom5/LGIII and Chrom4/LGVI flanked four and 74 candidate genes, respectively (Supplementary Table 3). Fifteen of the 74 genes were related to plant defense mechanisms. These included UDP formation and transportation, the integral component of membrane proteins, histone-lysine *N*-methyltransferase, phospholipid transport, actin cytoskeleton, calcium-ion binding, methyltransferase, UBIQUITIN-CONJUGATING ENZYME/RWD, AP-4 adaptor complex, oxidoreductase, acyl group transferases, hydrolases, G protein-coupled receptors, and protein involved in phosphorylation and proteolysis. Some of the genes were involved in pathways related to plant defense mechanisms. *Psat4g125440* is involved in cellulose biosynthesis, while *Psat4g111280*, *Psat4g110800*, *Psat4g108480*, and *Psat4g102720* are involved in protein ubiquitination.

DISCUSSION

Commercial farming in Canada is characterized by short rotations of cereal crops with canola and, to a limited extent, pulse crops. Disease surveys in Canada have identified *Fusarium* species as the most frequently isolated fungi from all crops surveyed for root rot severity (Chang, unpublished data). *Fusarium poae* was predominant in FHB-infected kernels, followed by *F. graminearum*; other *Fusarium* species were less common in infected kernels (Banik et al., 2019; Xue et al., 2019; Ziesman et al., 2019). The predominant *Fusarium* spp. isolated from the infected roots of field pea were *F. avenaceum*, *F. solani*, and *F. oxysporum* (Kraft, 1981; Kraft and Pflieger, 2001; Feng et al., 2010; Chittem et al., 2015; Rasiukevičiūtė et al., 2019). *Fusarium* species, especially *F. acuminatum*, have been reported to cause root rot of canola (Li et al., 2007; Chen et al., 2014).

Increasingly, *F. graminearum* has become a major problem across cereal-growing regions worldwide. For example, in

Manitoba, Canada, from 1937 to 1942, *F. graminearum* was present in <0.5% of 1,448, 262, 865, and 519 samples, respectively, of wheat, durum, barley, and oats tested, compared with 16.4–39.9% for *F. poae* and 13.5–29.5% for *F. acuminatum* (Gordon, 1944). In contrast, in Saskatchewan, Canada, from 2014 to 2018, *F. graminearum* represented 23.4–55.4% (mean, 39.1% over 5 years) of all the *Fusarium* species isolated from 1,812 wheat, 71 durum, 596 barley, and 177 oat samples (Olson et al., 2019). The increased frequency or shift to *F. graminearum* has also been reported in the US, China, Brazil, Argentina, Paraguay, Uruguay, and Africa (Savary et al., 2019). Unfortunately, damage to pulse crops by *F. graminearum* has not received enough attention compared with FHB of cereals. However, the available data suggest that, among pulse crops, field pea is most susceptible to *F. graminearum* (Clarkson, 1978; Chongo et al., 2001; Goswami et al., 2008; Bilgi et al., 2011; Foroud et al., 2014; Rasiukevičiūtė et al., 2019).

In a previous study, the pea cultivar “00-2067” was found to possess partial resistance to *Aphanomyces* root rot, while the cultivar “Reward” was susceptible (Wu et al., 2021). In this study, we screened the cultivars “00-2067” and “Reward” to determine their reaction to five isolates representing *F. solani*, *F. avenaceum*, *F. acuminatum*, *F. proliferatum*, and *F. graminearum*. The cultivar “00-2067” was partially resistant to all five species, which suggests that it might be tolerant to many pathogens of the pea root rot complex. The difference in disease severity between the mean root rot values of the two cultivars was significant ($p < 0.001$) only for the isolates representing *F. avenaceum* and *F. graminearum*. Therefore, the F₈ RIL population derived from “Reward” × “00-2067” was screened with F4A (*F. avenaceum*) and FG2 (*F. graminearum*) for the detection of partial resistance to the two *Fusarium* species. The greenhouse experiments were repeated four times to determine the G × E interaction for all traits. In addition, the best linear unbiased predictors (BLUPs) and LSM were applied to minimize environmental effects (Wang et al., 2018). The LSM identified six QTL, while BLUPs identified five QTL, suggesting that the LSM and BLUPs of the pooled data had comparable efficiency to detect important QTL.

Transgressive segregation was found for disease severity in the RILs inoculated with FG2 and F4A. This suggested that different resistance loci derived from the parental cultivars might have contributed to the stronger resistance observed in some of the RILs. Some transgressive RILs, such as X1303-19-3-1, X1303-21-3-1, X1303-26-2-1, X1304-21-3-1, and X1304-22-3-2, had lower disease severity in response to FG2 and higher vigor in all four environments compared with “00-2067.” In response to F4A, the RIL X1303-29-4-1 showed greater resistance and vigor compared with “00-2067.” Transgressive segregation was reported in other studies of resistance to *Fusarium* and *Aphanomyces* root rot in field pea (Feng et al., 2011; McPhee et al., 2012; Coyne et al., 2015, 2019; Nakedde et al., 2016; Wu et al., 2021). These transgressive lines will be valuable resources for developing commercial pea cultivars with improved resistance to *F. graminearum* and *F. avenaceum* and other pathogens of the pea root rot complex.

The average marker density of 1.8 marker/cM in this study was much greater than what has been reported in previous studies

of pea with PCR-based markers, while the total map length (1704.9 cM) was comparable. Feng et al. (2011) constructed a linkage map of 53 cM with 14 SSR markers and obtained a marker density of 0.26 marker/cM. McPhee et al. (2012) constructed a linkage map of total length 1,716 cM with 278 PCR-based markers and reported a marker density of 0.16 marker/cm. Similarly, Coyne et al. (2015) used 178 PCR-based markers to construct a linkage map of 1,323 cM and obtained a marker density of 0.13 marker/cM. More recently, Coyne et al. (2019) have applied 914 SNP markers to construct a linkage map of total length 1,073 cM and reported a marker density of 0.85 marker/cM. A marker density of 3.5 marker/cM and total map length of 843 cM were obtained when 18 pea lines were genotyped with the same SNP array set used in this study (Desgroux et al., 2016).

In this study, 11 QTL accounting for disease severity, vigor, and plant height were identified. The major QTL for disease resistance were located on Chrom4/LGIV, while two minor QTL were detected on Chrom5/LGIII and one QTL on Chrom3/LGV. These QTL were coincident with the QTL detected for resistance to *Aphanomyces* root rot (Wu et al., 2021). The major QTL ($R^2 = 68\text{--}80\%$) identified by McPhee et al. (2012) for resistance to *F. oxysporum* were also located on Chrom4/LGIV, while three minor QTL ($R^2 = 2.8\text{--}5.4\%$) were located on Chrom5/LGIII. Despite identifying the same chromosomes, the similarity of the location of the QTL cannot be confirmed, given the different markers used in the two studies. However, the coincidence of the QTL is not surprising, since very few partially resistant pea cultivars are used in breeding programs across the world. Feng et al. (2011) reported that the major QTL for root rot severity caused by *F. avenaceum* were located on Chrom7/LGVII. Coyne et al. (2015, 2019) reported that the major QTL for resistance to *F. solani* were located on Chrom6/LGII, while several minor QTL were located on Chrom5/LGIII, Chrom4/LGIV, Chrom6/LI, and Chrom7/LGVII.

Significant QTL \times QTL interactions were found between the minor QTL for disease severity and plant height but not for vigor. An interaction of the major QTL for disease severity, vigor, and height was not observed. Wu et al. (2021) reported that the same genomic regions controlled disease severity and vigor, while plant height was a poor measure of *Aphanomyces* root rot severity in pea. Coyne et al. (2019) treated plant height as a direct disease-related trait. In contrast, Desgroux et al. (2016) considered plant height as an agronomic trait. The reduced epistatic interaction might be due to a reduction in the detected number of additive-effect QTL from 27 in Wu et al. (2021) to 11 in the current study.

To the best of our knowledge, no genetic studies have been carried out to determine the genomic regions associated with the partial resistance of field pea to *F. graminearum*. The use of high-density SNP markers and SSR anchor markers contributed to the construction of a fine linkage map and the identification of two stable QTL located on Chrom4/LGIV associated with partial resistance to *F. graminearum*. The identified QTL showed broad resistance to *F. graminearum*, *F. solani*, *F. avenaceum*, *F. acuminatum*, and *F. proliferatum*, as well as to *A. euteiches*. This study, together with our previous report (Wu et al., 2021), suggests that “00-2067” and the transgressive RILs with lower

disease severity can be used to develop pea cultivars with improved root rot resistance.

DATA AVAILABILITY STATEMENT

Heritability of each trait, marker information, linkage information, and QTL profiles are available in the main manuscript or as **Supplementary Material**.

AUTHOR CONTRIBUTIONS

LW: inoculum preparation of *Fusarium* spp., greenhouse screening of RIL population and parents for resistance to *Fusarium* root rot, disease rating, measurement of vigor and plant height, phenotypic data analysis, DNA extraction, PCR, genotyping with SSR markers, and writing of the manuscript. RF-A: supervision of molecular marker work, molecular data analysis, linkage map construction, QTL mapping, and writing of the manuscript. SS: principal investigator, grant application, supervision and provision of technical support to graduate student, and revision of the manuscript. K-FC: grant application, supervision, and provision of technical support for graduate students for RIL population screening in the greenhouse. S-FH: principal investigator, grant application, supervision and provision of technical support for graduate students for RIL population screening in the greenhouse, and revision of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors thank the Canadian Agricultural Partnership (CAP) Program for providing financial support to this research under project no. 1000210132 and also acknowledge in-kind support from the University of Alberta and Alberta Agriculture and Forestry (Crop Diversification Centre – North), including access to greenhouse facilities and molecular laboratory equipment.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.784593/full#supplementary-material>

Supplementary Figure 1 | RIL population development from the parents: ‘Reward’ and ‘00-2067’.

Supplementary Table 1 | Root rot severity of the parental pea cultivars “00-2067” (partially resistant) and “Reward” (susceptible) to five *Fusarium* spp. under controlled conditions in the greenhouse.

Supplementary Table 2a | ANOVA for root rot severity, vigor, and plant height using the pooled data of an RIL population of pea inoculated with *Fusarium graminearum* isolate FG2 in four greenhouse experiments.

Supplementary Table 2b | ANOVA for root rot severity, vigor, and plant height using the pooled data of an RIL population of pea inoculated with *Fusarium avenaceum* isolate F4A in four greenhouse experiments.

Supplementary Table 3 | Candidate gene information for the major QTL.

REFERENCES

- Avenot, H., and Michailides, T. (2010). Progress in understanding molecular mechanisms and evolution of resistance to succinate dehydrogenase inhibiting (SDHI) fungicides in phytopathogenic fungi. *Crop Prot.* 29, 643–651.
- Bailey, K. L., Gossen, B. D., Gugel, R. K., and Morrall, R. A. A. (2003). *Diseases of Field Crops in Canada*, 3rd Edn. Saskatoon, SK: Canadian Phytopathological Society.
- Bailey, K. L., Mortensen, I. C., and Lafond, G. P. (1992). Effects of tillage systems and crop rotations on root and foliar diseases of wheat, flax, and peas in Saskatchewan. *Can. J. Plant Sci.* 22, 583–591. doi: 10.4141/cjps92-073
- Bainard, L. D., Navarro-Borrell, A., Hamel, C., Braun, K., Hanson, K., and Gan, Y. (2017). Increasing the frequency of pulses in crop rotations reduces soil fungal diversity and increases the proportion of fungal pathotrophs in a semiarid agroecosystem. *Agric. Ecosyst. Environ.* 240, 206–214.
- Banik, M., Beyene, M., and Wang, X. (2019). Fusarium head blight of barley in Manitoba – 2019. *Can. Plant Dis. Surv.* 100, 50–51.
- Bekkering, E. (2013). *Pulses in Canada*. CS96-325/2011-7E-PDF. Ottawa, ON: Statistics Canada.
- Bilgi, V. N., Bradley, C. A., Mathew, F. M., Ali, S., and Rasmussen, J. B. (2011). Root rot of dry edible bean caused by *Fusarium graminearum*. *Plant Health Prog.* 12:n1. doi: 10.1094/PHP-2011-0425-01-R
- Bing, D. J., Beauchesne, D., Turkington, T. K., Clayton, G., Sloan, A. G., Conner, R. L., et al. (2006). Registration of 'reward' field pea. *Crop Sci.* 46, 2705–2706.
- Brim, C. A. (1966). A modified pedigree method of selection in soybean. *Crop Sci.* 6:220. doi: 10.1007/BF00288795
- Chakraborty, S., Tiedemann, A. V., and Teng, P. S. (2000). Climate change: potential impact on plant diseases. *Environ. Pollut.* 108, 317–326. doi: 10.1016/S0269-7491(99)00210-9
- Chang, K. F. (unpublished data). *Root Rot Survey and Microorganism Isolation in Alberta in 2015–2020*.
- Chang, K. F., Bowness, R., Hwang, S. F., Turnbull, G. D., Howard, R. J., Lopetinsky, K., et al. (2005). Pea diseases occurring in central Alberta in 2004. *Can. Plant Dis. Surv.* 85, 89–90.
- Chang, K. F., Conner, R. L., Hwang, S. F., Ahmed, H. U., McLaren, D. L., Gossen, B. D., et al. (2014). Effects of seed treatment and inoculum density of *Fusarium avenaceum* and *Rhizoctonia solani* on seedling blight and root rot of faba bean. *Can. J. Plant Sci.* 94, 693–700. doi: 10.4141/cjps2013-339
- Chang, K. F., Hwang, S. F., Ahmed, H., Fu, H., Zhou, Q., Strelkov, S. E., et al. (2017). First report of *Phytophthora sansomeana* causing root rot in field pea in Alberta, Canada. *Crop Prot.* 101, 1–4. doi: 10.1016/j.cropro.2017.07.008
- Chang, K. F., Hwang, S. F., Ahmed, H., Gossen, B. D., Turnbull, G. D., and Strelkov, S. E. (2013). Management strategies to reduce losses caused by *Fusarium* seedling blight of field pea. *Can. J. Plant Sci.* 93, 619–625.
- Chang, K. F., Hwang, S. F., Ahmed, H. U., Strelkov, S. E., Harding, M. W., Conner, R. L., et al. (2018). Disease reaction to *Rhizoctonia solani* and yield losses in soybean. *Can. J. Plant Sci.* 98, 115–124.
- Chatterton, S., Bowness, R., and Harding, M. W. (2015). First report of root rot of field pea caused by *Aphanomyces euteiches* in Alberta, Canada. *Plant Dis.* 99, 288–289. doi: 10.1094/PDIS-09-14-0905-PDN
- Chatterton, S., Harding, M. W., Bowness, R., McLaren, D. L., Banniza, S., and Gossen, B. D. (2019). Importance and causal agents of root rot on field pea and lentil on the Canadian prairies, 2014–2017. *Can. J. Plant Pathol.* 41, 98–114. doi: 10.1080/07060661.2018.1547792
- Chen, Y., Zhou, Q., Strelkov, S. E., and Hwang, S. F. (2014). Genetic diversity and aggressiveness of *Fusarium* spp. isolated from canola in Alberta, Canada. *Plant Dis.* 98, 727–738. doi: 10.1094/PDIS-01-13-0061-RE
- Chittam, K., Mathew, F. M., Gregoire, M., Lamma, R. S., Chang, Y. W., Markell, S. G., et al. (2015). Identification and characterization of *Fusarium* spp. associated with root rots of field pea in North Dakota. *Eur. J. Plant Pathol.* 143, 641–649. doi: 10.1007/s10658-015-0714-8
- Chongo, G., Gossen, B. D., Kutcher, H. R., Gilbert, J., Turkington, T. K., Fernandez, M. R., et al. (2001). Reaction of seedling roots of 14 crop species to *Fusarium graminearum* from wheat heads. *Can. J. Plant Pathol.* 23, 132–137. doi: 10.1080/07060660109506920
- Cichy, K. A., Snapp, S., and Kirk, W. W. (2007). *Fusarium* root rot incidence and root system architecture in grafted common bean lines. *Plant Soil.* 300, 233–244. doi: 10.1007/s11104-007-9408-0
- Clarkson, J. D. S. (1978). Pathogenicity of *Fusarium* spp. associated with foot-rots of peas and beans. *Plant Pathol.* 27, 110–117.
- Conner, R. L., Chang, K. F., Hwang, S. F., Warkentin, T. D., and McRae, K. B. (2013). Assessment of tolerance for reducing yield losses in field pea caused by *Aphanomyces* root rot. *Can. J. Plant Sci.* 93, 473–482.
- Coyne, C., Pilet-Nayel, M., McGee, R., Porter, L., Smykal, P., and Grünwald, N. (2015). Identification of QTL controlling high levels of partial resistance to *Fusarium solani* f. sp. *pisi* in pea. *Plant Breed.* 134, 446–453.
- Coyne, C. J., Porter, L. D., Boutet, G., Ma, Y., McGee, R. J., Lesne, A., et al. (2019). Confirmation of *Fusarium* root rot resistance QTL Fsp-Ps 2.1 of pea under controlled conditions. *BMC Plant Biol.* 19:98. doi: 10.1186/s12870-019-1699-9
- Desgroux, A., L'Anthoëne, V., Roux-Duparque, M., Rivière, J., Aubert, G., Tayeh, N., et al. (2016). Genome-wide association mapping of partial resistance to *Aphanomyces euteiches* in pea. *BMC Genomics* 17:124. doi: 10.1186/s12864-016-2429-4
- Díaz Arias, M. M., Munkvold, G. P., and Leandro, L. F. (2011). First Report of *Fusarium proliferatum* causing root rot on soybean (*Glycine max*) in the United States. *Plant Dis.* 95:1316. doi: 10.1094/PDIS-04-11-0346
- Dorrance, A. E., McClure, S. A., and St Martin, S. K. (2003). Effect of partial resistance on *Phytophthora* stem rot incidence and yield of soybean in Ohio. *Plant Dis.* 87, 308–312. doi: 10.1094/PDIS.2003.87.3.308
- Elad, Y., and Pertot, I. (2014). Climate change impacts on plant pathogens and plant diseases. *J. Crop Improv.* 28, 99–139. doi: 10.1080/15427528.2014.865412
- Feng, J., Hwang, R., Chang, K. F., Conner, R. L., Hwang, S. F., Strelkov, S. E., et al. (2011). Identification of microsatellite markers linked to quantitative trait loci controlling resistance to *Fusarium* root rot in field pea. *Can. J. Plant Sci.* 91, 199–204.
- Feng, J., Hwang, R., Chang, K. F., Hwang, S. F., Strelkov, S. E., Gossen, B. D., et al. (2010). Genetic variation in *Fusarium avenaceum* causing root rot on field pea. *Plant Pathol.* 59, 413–421.
- Fernandez, M. R. (2007). *Fusarium* populations in roots of oilseed and pulse crops grown in eastern Saskatchewan. *Can. J. Plant Sci.* 87, 945–952.
- Fernandez, M. R., Huber, D., Basnyat, P., and Zentner, R. P. (2008). Impact of agronomic practices on populations of *Fusarium* and other fungi in cereal and noncereal crop residues on the Canadian prairies. *Soil Tillage Res.* 100, 60–71.
- Fletcher, J., Broadhurst, P., and Bansal, R. K. (1991). *Fusarium avenaceum*: a pathogen of lentil in New Zealand. *N. Zeal. J. Crop Hortic. Sci.* 19, 207–210.
- Foroud, N. A., Chatterton, S., Reid, L. M., Turkington, T. K., Tittlemier, S. A., and Gräfenhan, T. (2014). "Fusarium diseases of Canadian grain crops: impact and disease management strategies," in *Future Challenges in Crop Protection Against Fungal Pathogens*, eds A. Goyal and C. Manoharachary (New York, NY: Springer), 267–316. doi: 10.3389/fmicb.2018.01909
- Gautam, H. R., Bhardwaj, M., and Kumar, R. (2013). Climate change and its impact on plant diseases. *Curr. Sci.* 105, 1685–1691.
- Gill, K. (2018). Crop rotations compared to continuous canola and wheat for crop production and fertilizer use over six years. *Can. J. Plant Sci.* 98, 1139–1149. doi: 10.1139/CJPS-2017-0292
- Gordon, W. L. (1944). The occurrence of *Fusarium* species in Canada. I. species of *Fusarium* isolated from farm samples of cereal seed in Manitoba. *Can. J. Res.* 22, 282–286. doi: 10.1139/cjr44c-024
- Gossen, B. D., Conner, R. L., Chang, K. F., Pasche, J. S., McLaren, D. L., Henriquez, M. A., et al. (2016). Identifying and managing root rot of pulses on the Northern great plains. *Plant Dis.* 100, 1965–1978. doi: 10.1094/PDIS-02-16-0184-FE
- Goswami, R. S., Dong, Y., and Punja, Z. K. (2008). Host range and mycotoxin production by *Fusarium equiseti* isolates originating from ginseng fields1. *Can. J. Plant Pathol.* 30, 155–160. doi: 10.1080/07060660809507506
- Govaerts, B., Mezzalama, M., Unno, Y., Sayre, K. D., Luna-Guido, M., Vanherck, K., et al. (2007). Influence of tillage, residue management, and crop rotation on soil microbial biomass and catabolic diversity. *Appl. Soil Ecol.* 37, 18–30.
- Hwang, S. F., and Chang, K. F. (1989). Incidence and severity of root rot disease complex of field pea in northeastern Alberta in 1988. *Can. Plant Dis. Surv.* 69, 139–141.

- Hwang, S. F., Gossen, B. D., Turnbull, G. D., Chang, K. F., Howard, R. J., and Thomas, A. G. (2000). Effect of temperature, seeding date, fungicide seed treatment and inoculation with *Fusarium avenaceum* on seedling survival, root rot severity and yield of lentil. *Can. J. Plant. Sci.* 80, 899–907.
- Hwang, S. F., Howard, R. J., Chang, K. F., Park, B., and Burnett, P. A. (1994). Etiology and severity of *Fusarium* root rot of lentil in Alberta. *Can. J. Plant Pathol.* 16, 295–303. doi: 10.1080/07060669409500734
- Kaiser, W. J. (1992). Fungi associated with the seeds of commercial lentils from the US Pacific Northwest. *Plant Dis.* 76, 605–610. doi: 10.1094/pd-76-0605
- Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x
- Kraft, J. M. (1981). Registration of 792022 and 792024 pea germplasm (Reg Nos. GP21 and GP22). *Crop Sci.* 21, 352–353. doi: 10.2135/cropsci1981.001183x002100020047x
- Kraft, J. M., and Pfeiffer, F. L. (2001). *Compendium of Pea Diseases and Pests*. 2nd Edn. St. Paul, MN: The American Phytopathological Society.
- Kreplak, J., Madoui, M. A., Cápál, P., Novák, P., Labadie, K., Aubert, G., et al. (2019). A reference genome for pea provides insight into legume genome evolution. *Nat. Genet.* 51, 1411–1422. doi: 10.1038/s41588-019-0480-1
- Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69:101.
- Li, M., Gordon, M., and Gavin, A. (2007). New root diseases of canola in Australia. *Australas. Plant Dis. Notes* 2, 93–94. doi: 10.1071/DN07038
- Lincoln, S. M., Daly, M. J., and Lander, E. S. (1992). *Constructing genetic maps with MAPMAKER/EXP 3.0*. Whitehead Institute Technical Report. Cambridge, MA: Whitehead Institute.
- Loridon, K., McPhee, K., Morin, J., Dubreuil, P., Pilet-Nayel, M. L., Aubert, G., et al. (2005). Microsatellite marker polymorphism and mapping in pea (*Pisum sativum* L.). *Theor. Appl. Genet.* 111, 1022–1031. doi: 10.1007/s00122-005-0014-3
- McPhee, K. E., Inglis, D. A., Gundersen, B., and Coyne, C. J. (2012). Mapping QTL for *Fusarium* wilt race 2 partial resistance in pea (*Pisum sativum*). *Plant Breed.* 131, 300–306. doi: 10.1111/j.1439-0523.2011.01938.x
- 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 bi-parental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001
- Nakedde, T., Ibarra-Perez, F. J., Mukankusi, C., Waines, J. G., and Kelly, J. D. (2016). Mapping of QTL associated with *Fusarium* root rot resistance and root architecture traits in black beans. *Euphytica* 212, 51–63.
- Neumann, P., Pozárková, D., Vrána, J., Doležel, J., and Macas, J. (2002). Chromosome sorting and PCR-based physical mapping in pea (*Pisum sativum* L.). *Chromosome Res.* 10, 63–71. doi: 10.1023/a:1014274328269
- O'Donnell, K., Ward, T. J., Abern, D., Kistler, H. C., Aoki, T., Orwig, N., et al. (2008). Multilocus genotyping and molecular phylogenetics resolve a novel head blight pathogen within the *Fusarium graminearum* species complex from Ethiopia. *Fungal Genet. Biol.* 45, 1514–1522. doi: 10.1016/j.fgb.2008.09.002
- Olson, B., Blois, T., Ernst, B., Japp, M., Junek, S., Kutcher, H. R., et al. (2019). Seed-borne fusarium on cereal crops in Saskatchewan in 2018. *Can. Plant Dis. Surv.* 100, 100–103.
- Oyarzun, P. J. (1993). Bioassay to assess root rot in pea and effect of root rot on yield. *Neth. J. Plant Pathol.* 99, 61–75. doi: 10.1007/bf01998474
- Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., Zivot, E., et al. (2019). *Performance Analytics: Econometric Tools for Performance and Risk Analysis*. R package version. 1.5.3. Available online at: <https://CRAN.Rproject.org/package=PerformanceAnalytics> (accessed June 23, 2019).
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rasiukevičiūtė, N., Supronienė, S., Kelpsiene, J., Svegžda, P., Kadžienė, G., Šneideris, D., et al. (2019). Susceptibility of non-cereal crops to *Fusarium graminearum* complex and their role within cereal crop rotation as a source of inoculum for *Fusarium* head blight. *Span. J. Agric. Res.* 16:1012.
- Safarieskandari, S., Chatterton, S., and Hall, L. (2020). Pathogenicity and host range of *Fusarium* species associated with pea root rot in Alberta, Canada. *Can. J. Plant Pathol.* 43, 162–171. doi: 10.1080/07060661.2020.1730442
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* 3, 430–439. doi: 10.1038/s41559-018-0793-y
- Schneider, K. A., Grafton, K. F., and Kelly, J. D. (2001). QTL analysis of resistance to *Fusarium* root rot in Bean. *Crop Sci.* 41, 535–542. doi: 10.2135/cropsci2001.412535x
- Schwartz, H. F., Steadman, J. R., Hall, R., and Forster, R. L. (2005). *Compendium of Bean Diseases*, 2nd Edn. St. Paul, MN: APS Press, 109.
- Son, H., Kim, M. G., Min, M., Seo, Y. S., Lim, J. Y., Choi, G. J., et al. (2013). AbaA regulates conidiogenesis in the ascomycete fungus *Fusarium graminearum*. *PLoS One* 8:e72915. doi: 10.1371/journal.pone.0072915
- Su, Y., Gabrielle, B., Beillouin, D., and Makowski, D. (2021). High probability of yield gain through conservation agriculture in dry regions for major staple crops. *Sci. Rep.* 11:3344. doi: 10.1038/s41598-021-82375-1
- Tayeh, N., Aluome, C., Falque, M., Jacquin, F., Klein, A., Chauveau, A., et al. (2015). Development of two major resources for pea genomics: the GenoPea 13.2K SNP array and a high-density, high-resolution consensus genetic map. *Plant J.* 84, 1257–1273. doi: 10.1111/tpj.13070
- Tyler, B. M. (2007). *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Mol. Plant Pathol.* 8, 1–8. doi: 10.1111/j.1364-3703.2006.00373.x
- van der Lee, T., Zhang, H., van Diepeningen, A., and Waalwijk, C. (2015). Biogeography of *Fusarium graminearum* species complex and chemotypes: a review. *Food Addit. Contam.* 32, 453–463. doi: 10.1080/19440049.2014.984244
- 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
- Wang, S., Basten, C. J., and Zeng, Z. B. (2012). *Windows QTL Cartographer 2.5*. Raleigh, NC: Department of Statistics.
- Wang, Y., VandenLangenberg, K., Wen, C., Wehner, T. C., and Weng, Y. (2018). QTL mapping of downy and powdery mildew resistances in PI 197088 cucumber with genotyping-by-sequencing in RIL population. *Theor. Appl. Genet.* 131, 597–611. doi: 10.1007/s00122-017-3022-1
- Wille, L., Messmer, M. M., Bodenhausen, N., Studer, B., and Hohmann, P. (2020). Heritable variation in pea for resistance against a root rot complex and its characterization by amplicon sequencing. *Front. Plant Sci.* 11:542153. doi: 10.3389/fpls.2020.542153
- Wu, L., Chang, K., Hwang, S., Conner, R., Fredua-Agyeman, R., Feindel, D., et al. (2019). Evaluation of host resistance and fungicide application as tools for the management of root rot of field pea caused by *Aphanomyces euteiches*. *Crop J.* 7, 38–48. doi: 10.1016/j.cj.2018.07.005
- Wu, L., Fredua-Agyeman, R., Hwang, S. F., Chang, K. F., Conner, R. L., McLaren, D. L., et al. (2021). Mapping QTL associated with partial resistance to *Aphanomyces* root rot in pea (*Pisum sativum* L.) using a 13.2 K SNP array and SSR markers. *Theor. Appl. Genet.* 134, 2965–2990. doi: 10.1007/s00122-021-03871-6
- Wu, L. F., Chang, K. F., Fu, H., Akter, I., Li, N., Hwang, S. F., et al. (2017). The occurrence of and microorganisms associated with root rot of field pea in Alberta in 2016. *Can. Plant Dis. Surv.* 97, 193–195.
- Wu, Y., Bhat, P. R., Close, T. J., and Lonardi, S. (2008). Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet.* 4:e1000212. doi: 10.1371/journal.pgen.1000212
- Xue, A. G., Chen, Y., Seifert, K., Guo, W., Blackwell, B. A., Harris, L. J., et al. (2019). Prevalence of *Fusarium* species causing head blight of spring wheat, barley and oat in Ontario during 2001–2017. *Can. J. Plant Pathol.* 41, 392–402. doi: 10.1080/07060661.2019.1582560
- Xue, A. G., Tuey, H. J., and Mathur, S. (1998). Diseases of field pea in Manitoba in 1997. *Can. Plant Dis. Surv.* 78, 97–98.
- Xue, A. G., Tuey, H. J., and Platford, R. G. (2000). Diseases of field pea in Manitoba in 1999. *Can. Plant Dis. Surv.* 80, 110–111.
- Zhao, P. (2020). *A Tool for Phenotypic Data Processing*. R package version 0.1.0. Available online at: <https://CRAN.Rproject.org/package=Phenotype> (accessed July 31, 2020).
- Zhou, Q., Chen, Y., Yang, Y., Ahmed, H., Hwang, S. F., and Strelkov, S. E. (2014). Effect of inoculum density and quantitative PCR-based detection of *Rhizoctonia solani* AG-2-1 and *Fusarium avenaceum* on canola. *Crop Prot.* 59, 71–77.
- Ziesman, B., Bawolin, C., Sliva, T., Hartley, S., Hladun, M., Fernandez, M. R., et al. (2019). *Fusarium* head blight in barley in Saskatchewan 2018. *Can. Plant Dis. Surv.* 99, 69–70.

- Zitnick-Anderson, K., del Río Mendoza, L. E., Forster, S., and Pasche, J. S. (2020). Associations among the communities of soil-borne pathogens, soil edaphic properties and disease incidence in the field pea root rot complex. *Plant Soil* 457, 339–354. doi: 10.1007/s11104-020-04745-4
- Zitnick-Anderson, K., Simons, K., and Pasche, J. S. (2018). Detection and qPCR quantification of seven *Fusarium* species associated with the root rot complex in field pea. *Can. J. Plant Pathol.* 40, 261–271. doi: 10.1080/07060661.2018.1429494

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 Wu, Fredua-Agyeman, Strelkov, Chang and Hwang. 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.



Quantitative Trait Locus Mapping for Resistance Against *Pyrenopeziza brassicae* Derived From a *Brassica napus* Secondary Gene Pool

Chinthani S. Karandeni Dewage^{1*}, Katherine Cools^{2†}, Henrik U. Stotz¹, Aiming Qi¹, Yong-Ju Huang¹, Rachel Wells³ and Bruce D. L. Fitt¹

¹ Centre for Agriculture, Food, and Environmental Management Research, School of Life and Medical Sciences, University of Hertfordshire, Hatfield, United Kingdom, ² Rothamsted Research, Harpenden, United Kingdom, ³ John Innes Centre, Norwich Research Park, Norwich, United Kingdom

OPEN ACCESS

Edited by:

Anna Maria Mastrangelo,
Research Centre for Cereal
and Industrial Crops, Council
for Agricultural and Economics
Research (CREA), Italy

Reviewed by:

Javier Sánchez-Martín,
University of Zurich, Switzerland
Zhansheng Li,
Institute of Vegetables and Flowers,
Chinese Academy of Agricultural
Sciences (CAAS), China

*Correspondence:

Chinthani S. Karandeni Dewage
c.s.karandeni-dewage@herts.ac.uk

† Present address:

Katherine Cools,
Postharvest Bioscience Consultant,
Binfield, United Kingdom

Specialty section:

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

Received: 29 September 2021

Accepted: 03 January 2022

Published: 04 February 2022

Citation:

Karandeni Dewage CS, Cools K,
Stotz HU, Qi A, Huang YJ, Wells R
and Fitt BDL (2022) Quantitative Trait
Locus Mapping for Resistance
Against *Pyrenopeziza brassicae*
Derived From a *Brassica napus*
Secondary Gene Pool.
Front. Plant Sci. 13:786189.
doi: 10.3389/fpls.2022.786189

Use of host resistance is the most economical and environmentally safe way to control light leaf spot disease of oilseed rape (*Brassica napus*). The causal organism of light leaf spot, *Pyrenopeziza brassicae*, is one of the most economically damaging pathogens of oilseed rape in the United Kingdom and it is considered to have a high potential to evolve due to its mixed reproduction system and airborne ascospores. This necessitates diverse sources of host resistance, which are inadequate at present to minimize yield losses caused by this disease. To address this, we screened a doubled haploid (DH) population of oilseed rape, derived from a secondary gene pool (ancestral genomes) of *B. napus* for the introgression of resistance against *P. brassicae*. DH lines were phenotyped using controlled-environment and glasshouse experiments with *P. brassicae* populations obtained from three different geographic locations in the United Kingdom. Selected DH lines with different levels of resistance were further studied in a controlled-environment experiment using both visual (scanning electron microscope – SEM) and molecular (quantitative PCR) assessment methods to understand the mode/s of host resistance. There was a clear phenotypic variation for resistance against *P. brassicae* in this DH population. Quantitative trait locus (QTL) analysis identified four QTLs with moderate to large effects, which were located on linkage groups C1, C6, and C9. Of these, the QTL on the linkage group C1 appeared to have a major effect on limiting *P. brassicae* asexual sporulation. Study of the sub-cuticular growth phase of *P. brassicae* using qPCR and SEM showed that the pathogen was able to infect and colonise both resistant and susceptible Q DH lines and control *B. napus* cultivars. However, the rate of increase of pathogen biomass was significantly smaller in resistant lines, suggesting that the resistance segregating in this DH population limits colonisation/sporulation by the pathogen rather than eliminating the pathogen. Resistance QTLs identified in this study provide a useful resource for breeding cultivar resistance for effective control of light leaf spot and form a starting point for functional identification of the genes controlling resistance against *P. brassicae* that can contribute to our knowledge on mechanisms of partial resistance of crops against pathogens.

Keywords: crop losses, fungal pathogens, host resistance, light leaf spot, oilseed rape, quantitative resistance

INTRODUCTION

Host plant resistance against pathogens is an important characteristic in agricultural crops. In general, resistance against pathogens is described under two broad categories: complete/qualitative resistance and incomplete/quantitative resistance (Roux et al., 2014; French et al., 2016). Of these two categories, quantitative disease resistance (QDR) is preferred as a broad-spectrum, durable source of resistance (Poland et al., 2009; French et al., 2016). Development of genetic linkage maps with DNA-based markers and quantitative trait loci (QTLs) mapping by exploiting the marker-trait associations have proved to be effective in breeding for plant disease resistance (Poland et al., 2009; St. Clair, 2010). Further dissection of resistance loci and characterisation of underlying genes can improve understanding of the mechanisms of QDR against plant pathogens (French et al., 2016), especially for those with complex host-pathogen interactions (e.g., extracellular pathogens) where the operation of host resistance does not eliminate the pathogen.

Light leaf spot (LLS), caused by ascomycete extracellular fungal pathogen *Pyrenopeziza brassicae* Sutton and Rawlinson (anamorph *Cylindrosporium concentricum* Grev.) (Rawlinson et al., 1978) is one of the most widespread diseases of oilseed rape (*Brassica napus* L.) in the United Kingdom, causing major yield losses (CropMonitor, 2016). The pathogen affects most aerial parts of the plant, including leaves, stems, flowers, and seed pods, resulting in reduction of leaf photosynthetic area, reduced plant vigour, and further yield loss through pod shatter (Boys et al., 2007). According to recent reports, *P. brassicae* can cause up to 30% yield reduction (AHDB, 2021). However, earlier studies have reported yield losses as great as 50% under severe epidemics (Rawlinson et al., 1978). Current recommendations for managing LLS risks include application of fungicides, growing cultivars with good field resistance, and crop sanitary practices, such as ploughing crop debris, delaying of the sowing date, and separation of oilseed rape crops in space and time (AHDB, 2021). In practice, the effectiveness of fungicide control methods depends on several factors, including fungicide application timings, weather, and shifts in *P. brassicae* populations toward fungicide insensitivity. Additionally, fungicide applications may not always be an economically viable solution for farmers. Despite crop sanitary practices, cruciferous vegetables, weeds species, and volunteer oilseed rape plants that occur from pod shatter during harvest can provide a pathway for the pathogen to transfer between cropping seasons due to the cross-infectivity of *P. brassicae* between oilseed rape and other *Brassica* species (Maddock et al., 1981; Evans et al., 2003). Therefore, it is necessary to put more emphasis on host resistance in LLS management practices and produce cultivars that have a greater economic return to sustain the production of oilseed rape, which is the third largest arable crop in the United Kingdom (DEFRA, 2020). The development of oilseed rape cultivars with good levels of field resistance against *P. brassicae* can provide economical means of disease control, especially for farmers with small to medium-sized arable farming areas.

Even though the average LLS resistance rating of oilseed rape cultivars has increased in recent years¹, frequent recent epidemics of LLS indicate that the currently available cultivar resistance is inadequate to achieve successful control of this disease (CropMonitor, 2016). There has been little knowledge on the genetic basis of host resistance against *P. brassicae* (Karandeni Dewage et al., 2021) with only three published studies on mapping qualitative or quantitative resistance genes (Pilet et al., 1998; Bradburne et al., 1999; Boys et al., 2012). Moreover, sexual reproduction of *P. brassicae* could lead to the development of new virulent strains, rendering the resistance genes ineffective (Boys et al., 2007; Karandeni Dewage et al., 2018). Therefore, diversification of the resistance sources is essential to achieve effective and prolonged control of LLS through host resistance. Improvement of cultivar resistance through selective breeding requires sufficient genetic diversity to be present within the current gene pool. Nevertheless, continuous selection of plant material for specific traits can cause genetic bottlenecks, resulting in reduced genetic diversity in the primary gene pool (Doebley et al., 2006). Narrow gene pools can make crop species more vulnerable to emerging pests and pathogens and reduce the potential for improving crop productivity (Hyten et al., 2006). In such cases, genetic variations present in external gene pools provide plant breeders with an opportunity to improve crop cultivars by incorporating various traits for which there is insufficient diversity in the primary gene pools (Boyd et al., 2013).

Oilseed rape has been known for its narrow genetic diversity caused by strong selection for various traits in its breeding history (Snowdon and Iniguez Luy, 2012). For example, the process of developing modern double-low cultivars (low seed glucosinolate and low erucic acid contents) is considered to have caused genetic bottlenecks in current *B. napus* gene pools. However, being a natural hybrid between *B. rapa* and *B. oleracea* (Chalhoub et al., 2014), the compatibility between these two ancestral species (secondary gene pool) of *B. napus* enables introgression of new sources of resistance from these external gene pools into cultivated *B. napus*. Compared to *B. napus*, *B. rapa*, and *B. oleracea* are considered to have higher genetic diversity and they have proved to be effective in providing new resistance genes against other important pathogens of oilseed rape (Neik et al., 2017; Robin et al., 2017; Katche et al., 2019). Experimental work described in this article has focused on analysing a doubled haploid (DH) population of oilseed rape, derived from the ancestral genomes of *B. napus*, as a potential source of resistance against the LLS pathogen *P. brassicae* and understanding the operation of host resistance against this pathogen.

MATERIALS AND METHODS

Plant Material

A progeny of DH lines (Q population consisting of a total of 92 lines) derived from an F1 cross between synthetic *B. napus* (*B. oleracea atlantica* X *B. rapa oleifera* “29”) and oilseed rape cv.

¹<https://ahdb.org.uk/rlarchive>

Tapidor (European winter-type cultivar with a strict vernalisation requirement and low seed erucic acid and low glucosinolate content) (Mithen and Magrath, 1992; Smooker et al., 2011) was used in this study. Additionally, oilseed rape cultivars Canberra [UK Agriculture and Horticulture Development Board (AHDB) recommended list (RL) resistance rating 7 (2007/08), resistant], Cuillin [RL resistance rating 8 (2014/15), resistant], Marathon [RL resistance rating 5 (2016/17), susceptible], Bristol [RL resistance rating 2 (1996/97), susceptible], Imola (characteristic black flecking resistance phenotype against *P. brassicae* infection, Boys et al., 2012), and Tapidor and *B. rapa oleifera* “29” (A-genome parent of synthetic *B. napus*) were included in the phenotyping experiments. The AHDB RL ratings (resistance rating for LLS on 1–9 scale, where nine is most resistant), which are available for commercial cultivars, were taken from the most recent records available (given in parentheses next to the RL rating).

Phenotyping of Resistance Against *Pyrenopeziza brassicae* in the Q Doubled Haploid Population

The Q DH population was assessed for its resistance against *P. brassicae* in three separate experiments to represent different *P. brassicae* populations (mixture of isolates collected from diseased leaves from oilseed rape crops) and different environmental conditions. These consisted of two glasshouse experiments (GH) and one controlled-environment (CE) experiment. Each experiment included appropriate resistant/susceptible control cultivars. The numbers of Q DH lines and resistant/susceptible control cultivars included in each of the three phenotyping experiments are given in **Table 1**. Different *P. brassicae* populations that originated from England or Scotland were used in the three experiments (**Table 1**). *Pyrenopeziza brassicae* conidial suspensions were prepared by selecting diseased oilseed rape leaves with clear LLS symptoms and incubating them at 4°C for 5 days in sealed polyethylene bags with a layer of dampened paper towel to increase humidity to induce sporulation. Leaves were then washed with sterile distilled water to produce conidial inoculum. Conidial suspensions were filtered through sterile Miracloth (Calbiochem, United States), and the concentration of each spore suspension was measured using a haemocytometer. Spore concentration was adjusted with sterile distilled water to 10⁵ spores/ml for the glasshouse experiments and 10⁴ spores/ml for the CE experiment and suspensions were stored at –20°C until needed.

Glasshouse experiments were arranged in an alpha design generated using an alpha design generator (Parsad et al., 2007) as it was not possible to assess all the lines/cvs in one experiment due to space limitations. Q DH lines were divided into small batches and assessed at different occasions within each glasshouse experiment (four and three batches in the first and second glasshouse experiment, respectively). Resistant/susceptible control cultivars and a Q DH line (Q12) were repeated on each occasion in GH1 and GH2 to monitor the uniformity of experimental conditions. Plants were grown

in 9 cm diameter pots until they reached growth stage 1.4–1.5 (plants have five true leaves) (Sylvester-Bradley et al., 1984). Five replicate plants were included for each Q DH line and control cultivar. At growth stages 1.4–1.5, plants were spray-inoculated with *P. brassicae* conidial suspensions using a 50 ml travel spray bottle (Boots, United Kingdom) until all the leaves were evenly and fully covered with fine droplets of the spore suspension. After inoculation, plants were covered with a polyethylene cover for 48 h to maintain high humidity to facilitate spore germination and infection. Glasshouse conditions were maintained at 16°C/14°C day/night temperatures with natural daylight and supplemented by a 12 h photoperiod. At 24 days post-inoculation (dpi), plants were destructively harvested by cutting at the stem base above the compost surface, individually placed in polyethylene bags with a dampened paper towel, and incubated at 4°C for 5 days to induce sporulation. The disease assessment was made by visually estimating the percentage leaf area covered with *P. brassicae* acervuli (sporulation). The presence or absence of a necrotic response was also recorded for each plant.

In the CE experiment, plants were first grown in 7 cm diameter pots and maintained in a glasshouse at 20°C for 3 weeks. Four replicate plants were included for each Q DH line and the control cultivar. At growth stages 1.4–1.5, plants were spray-inoculated using an aerosol sprayer (Chrom Atomiser, Camlab; Cambridge, United Kingdom) until drops ran off the leaves. After inoculation, plants were individually covered with polyethylene bags (26 cm × 38 cm) to maintain high humidity to facilitate spore germination and infection and kept in a CE room at 16°C with a 12 h photoperiod. Bags were removed 48 h after inoculation and were replaced 14 days later for the final week before disease assessment. LLS severity on each plant was assessed by visually estimating the percentage leaf area covered with *P. brassicae* acervuli.

Statistical Analysis and Mapping of Quantitative Trait Locus for Resistance Against *Pyrenopeziza brassicae*

Light leaf spot severity data (% leaf area covered with *P. brassicae* sporulation) were subjected to ANOVA using GENSTAT statistical software for Windows (Payne et al., 2011). Arcsine transformation of the percentage of leaf area with sporulation was made with an arcsine formula in Excel before ANOVA was done, so that the variance was more homogeneous across treatments and measurements were normally distributed. For each glasshouse experiment, data generated by the alpha design experiments were combined and the effect of experiment was analysed as a factor in the ANOVA using control cultivars and Q DH lines, which were included in each experiment (batch). For analysis of the relationships between LLS severities measured in each of the three phenotyping experiments, simple linear regression analyses were done using calculated means for different Q DH lines and cultivars.

The linkage map and marker data for the Q DH population have been described previously (Smooker et al., 2011). The genetic linkage map of the Q DH population comprised of

TABLE 1 | Summary of phenotyping experiments used to assess resistance against *Pyrenopeziza brassicae* in the *Brassica napus* Q doubled haploid (DH) population.

| Experiment | Number of Q DH lines ^a | Control cultivars ^b | Origin of <i>Pyrenopeziza brassicae</i> inoculum ^c | Inoculum concentration (spores/ml) | Light leaf spot severity (% leaf area affected) (range) |
|-----------------------------|-----------------------------------|--|---|------------------------------------|---|
| Glasshouse 1 (GH1) | 84 | Marathon (S), Bristol (S), Cuillin (R), Imola (R), Tapidor, <i>Brassica rapa oleifera</i> "29" | Morley, Norfolk, England | 1 × 10 ⁵ | 0–83% |
| Glasshouse 2 (GH2) | 78 | Bristol (S), Charger (S), Imola (R) | Aberdeen, Scotland | 1 × 10 ⁵ | 0–63% |
| Controlled environment (CE) | 89 | Canberra (R) | Harpenden, Hertfordshire, England | 1 × 10 ⁴ | 0–46% |

^aQ DH population consists of 92 lines in total. From GH1 and GH2, 77 and 70 lines were taken into the final data analysis, respectively. Therefore, there were 70 lines in common between the three experiments.

^bEach experiment included resistant (R)/susceptible (S) control cultivars. Additionally, GH1 experiment included two of the parental lines of the Q DH population; *B. rapa oleifera* "29" (A-genome parent of the synthetic *B. napus*) and cv. Tapidor.

^cPlants were inoculated with *P. brassicae* populations (conidial suspensions collected from diseased leaves from oilseed rape crops in England or Scotland).

358 simple sequence repeats (SSR) markers over 19 linkage groups with a total genetic distance of 1,381 cm. QTL mapping was implemented with QTL cartographer version 2.5 (Wang et al., 2012) using quantitative disease severity (% leaf area with *P. brassicae* sporulation) data within each individual experiment and across all three experiments using combined data. Initial genome scans for marker-trait associations were done using single marker analysis to identify possible QTLs. The results obtained from single marker analysis were further refined using interval mapping (IM) (Haley and Knott, 1992). Genome-wide QTL threshold was determined by permutation analysis using 1,000 iterations corresponding to a significance level of $\alpha = 0.10$. Support interval for each QTL was determined based on the decrease in 1.5-logarithm of the odds (LOD) on either side of the LOD maximum (Silva et al., 2012). Since IM may be affected by the skewed distribution of phenotype data, QTL positions and the effects were confirmed with transformed data. The binary (categorical) phenotype data for necrosis obtained from the two glasshouse experiments (presence or absence of a necrotic response) were analysed using the non-parametric Kruskal–Wallis test. QTLs detected by IM were visualised on the linkage map of the Q DH population using MapChart software (version 2.32) (Voorrips, 2002) with manual editing.

Assessment of the Sub-Cuticular Growth Phase of *Pyrenopeziza brassicae* in Q Doubled Haploid Lines

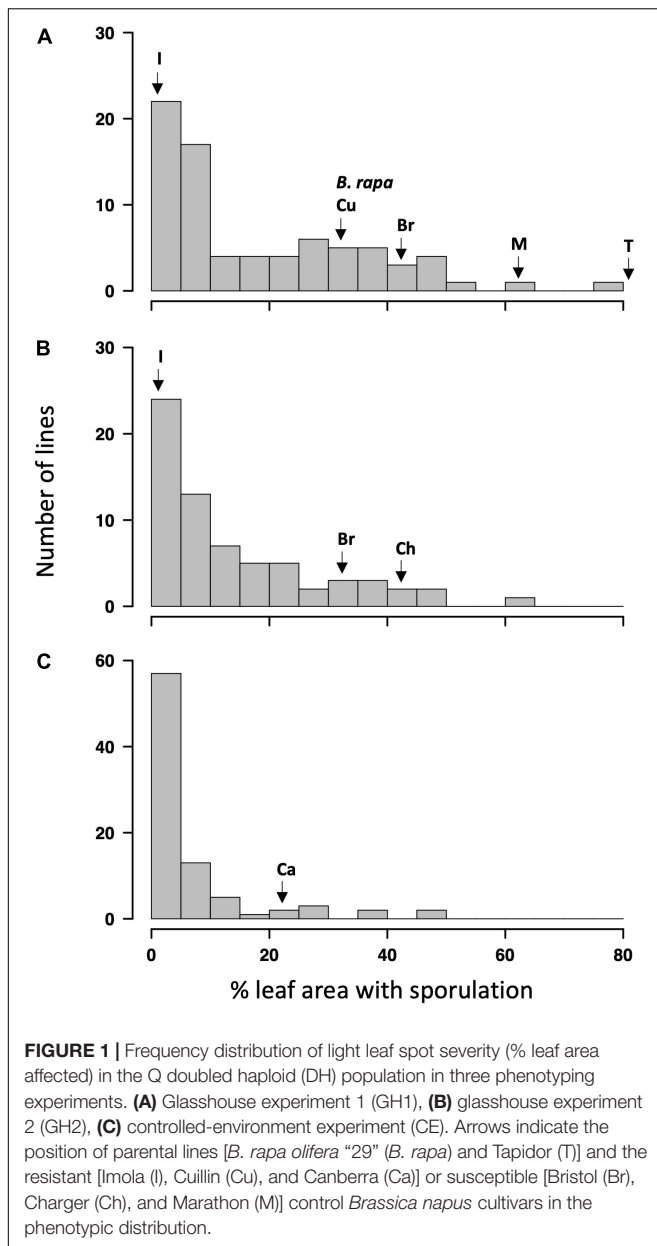
Four Q DH lines (Q04, Q38, Q69, and Q83), based on the amount of *P. brassicae* sporulation and presence of a necrotic response observed in phenotyping experiments, and oilseed rape cultivars Bristol and Imola were selected. Plants were grown in 9 cm diameter pots under controlled-environment conditions (FITOCLIMA D1200, ARALAB, Rio de Mouro, Portugal) with a 12 h photoperiod, 60% relative humidity, and 20°C/18°C day/night temperatures, respectively. Plants were arranged in a randomised complete block design generated using Experimental Design Generator and Randomiser (EDGAR) (Brown, 2004). At growth stage 1.4–1.5 (five leaves unfolded) (Sylvester-Bradley et al., 1984), plants were point-inoculated at four "marked" points on adaxial surfaces of each of the fourth and fifth true leaves using sterilised Whatman no. 1 filter papers (cut

into c.0.8 mm × 0.8 mm squares) immersed in a *P. brassicae* conidial suspension (10⁵ spores/ml). After inoculation, plants were covered with a polyethylene cover for 48 h to maintain high humidity. Inoculated plants were maintained in the controlled environment cabinet with a 12 h photoperiod, 60% relative humidity, and 16°C/14°C day/night temperatures, respectively.

Plants were sampled at 0, 3, 7, 14, and 24 dpi, and the fourth and fifth true leaves were used for the analysis of sub-cuticular growth of *P. brassicae* using quantitative PCR (qPCR) with species-specific primers (Karolewski et al., 2006) and scanning electron microscopy (SEM), respectively. For qPCR analysis, the fourth true leaf to appear was removed from each plant and 2 cm diameter leaf discs were cut from each inoculation point. Leaf discs were individually placed in 2 ml tubes, frozen at –20°C, and freeze-dried. Samples were processed in a Fastprep machine (MP Biomedicals, United Kingdom) with three metal beads (3 mm diameter) until leaf discs were ground to a fine powder. DNA extraction and quantification of *P. brassicae* DNA was done according to the method described in Boys et al. (2012) with minor modifications (DNA samples were diluted to a final concentration of 20 ng/μl and five standards, each containing known quantities of *P. brassicae* DNA ranging from 1 pg to 10 ng, were used in qPCR). Quantitative PCR data were analysed by simple linear regression of *P. brassicae* DNA content against the days post-inoculation (dpi). Since the amount of *P. brassicae* DNA in leaf tissues showed an exponential increase with time after inoculation, data were transformed by taking the common logarithm (log₁₀) of the original measurements. A position and parallelism regression analysis was used to analyse the differences in the increase in the amount of *P. brassicae* DNA over time between the six lines/cultivars included in this experiment. All the analyses were done using GENSTAT statistical software for Windows (Payne et al., 2011).

Scanning electron microscope (SEM) analysis of leaf samples of the Q DH lines collected at different time points after inoculation was done using the Bioimaging facility at Rothamsted Research, Harpenden, United Kingdom.² Pieces of leaves (c. 4 mm × 4 mm) were obtained from the inoculation points using a sterile blade and were prepared according to standard operating procedures of the SEM instrument (JEOL JSM-6360, JEOL Ltd.,

²<https://www.rothamsted.ac.uk/bioimaging>



United Kingdom) at Rothamsted Research for examination and recording images.

RESULTS

Phenotyping of Resistance Against *Pyrenopeziza brassicae* in the Q Doubled Haploid Population

The ANOVA indicated a significant effect of genotypes ($p < 0.01$) on the LLS severity. The effect of the batch experiment was not significant within GH1 and GH2 ($p > 0.69$, **Supplementary Table 1**). Distribution of resistance against *P. brassicae*, measured using % leaf area covered with pathogen sporulation, among

Q DH lines is illustrated in **Figure 1**. The phenotype distribution showed positive skew toward resistance in all three experiments. Significant positive correlations were observed between the experiments (**Supplementary Table 2**; $p < 0.01$), indicating consistency in disease scoring methodology across experiments. The correlation was particularly good between the two glasshouse experiments, and the overall LLS severity in these two experiments appeared to be greater [ranging from 0 to 83% and from 0 to 63% leaf area affected in the first (GH1) and second (GH2) glasshouse experiments, respectively] compared to that of the CE experiment (ranging from 0 to 46%). Some of the Q DH lines had disease severities smaller than or equal to those of the resistant control cultivars (cvs Imola and Cuillin) in GH1 and GH2 experiments, and most of the Q DH lines had disease severities smaller/equal to those of the resistant control (cv. Canberra) in the CE experiment. Thirty-nine lines in GH1 and 27 lines in GH2 showed no significant difference from the disease severity observed on cv. Imola (0.9 and 0% leaf area with sporulation in GH1 and GH2, respectively) ($p < 0.05$). The two parental lines included in the GH1 experiment differed significantly in LLS severity ($p < 0.05$). Of these, the A-genome parent of the synthetic *B. napus*, *B. rapa oleifera* “29” showed an average of 32.5% leaf area with sporulation, which is similar to that of cv. Cuillin. In contrast, cv. Tapidor showed extreme susceptibility to *P. brassicae* with an average of 83% leaf area with sporulation.

In addition to the varying numbers of *P. brassicae* acervuli that appeared with or without lesion formation, some of the Q DH lines showed a necrotic response against *P. brassicae* that started to appear c. 10–14 dpi (**Figures 2A–C**). These responses were mainly observed on the leaf veins, midribs, and along petioles, and elsewhere on the leaf lamina (leaf blade). This was similar to the necrosis observed on cv. Imola (**Figure 2D**) that is known to contain a major-gene locus for resistance against *P. brassicae*. However, cv. Imola was consistent in producing zero to very little ($>1\%$ area affected) sporulation (mainly confined to leaf veins, the midrib, and petioles) in different experiments, whereas the Q DH lines with necrosis showed a great variation in sporulation. For example, some of the Q DH lines appeared to have large numbers of acervuli in the presence of a less intense necrotic response. Some lines showed limited sporulation confined only to the leaf veins and the midribs (**Figure 2E**). Necrotic flecking on the leaf lamina appeared in concentric rings (**Figure 2F**) which resembled concentric ring-like sporulation patterns characteristic of susceptible interactions (**Figure 2G**). In the GH1 experiment, 41 out of 77 lines produced a necrotic response, whereas in GH2 experiment, 46 out of 70 lines produced necrosis. Of these, six lines that had necrosis in GH1 did not produce necrosis in GH2, and nine lines that did not have necrosis in GH1 produced necrosis in GH2. Comparisons of the LLS severities between Q DH lines with or without a necrotic response using qualitative assessments (presence or absence) made in GH1 and GH2 experiments are illustrated in **Figure 3**. According to Shapiro-Wilk W statistics, LLS severity distributions showed deviation from normality ($p < 0.05$) in both groups. The difference in the median values of % leaf area covered with acervuli between Q

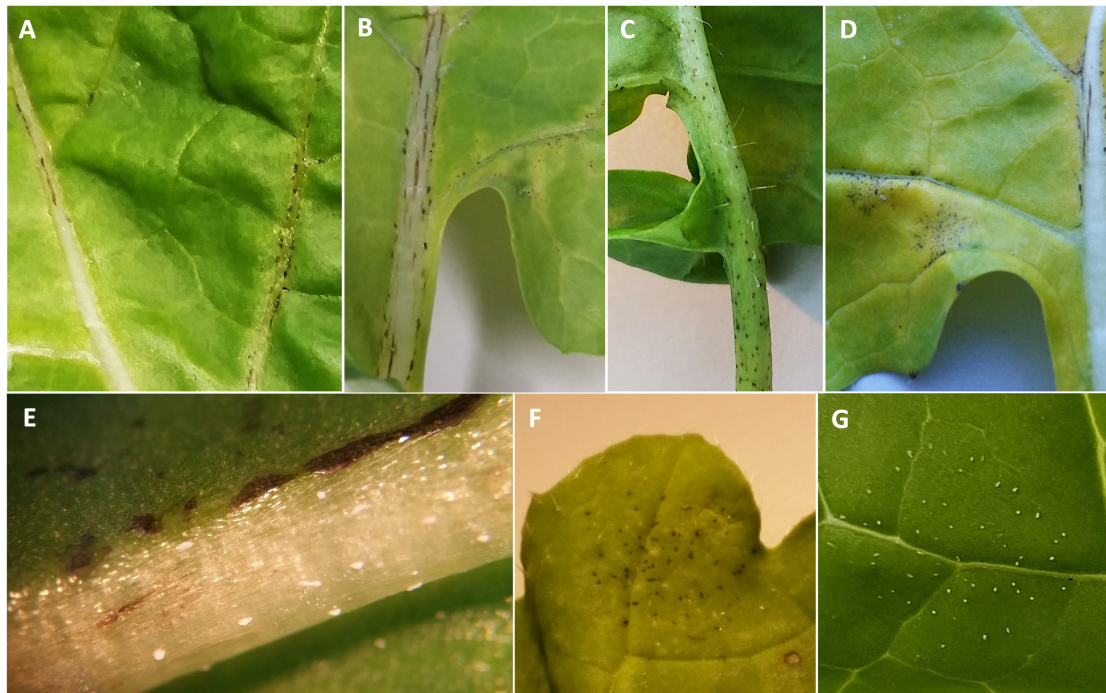


FIGURE 2 | Necrotic responses observed in different *B. napus* lines from the Q DH population. **(A)** Q88, **(B)** Q60, **(C)** Q83, **(D)** cv. Imola, **(E)** Q64 with necrosis and occasional acervuli on the midrib, **(F)** Q33 with necrosis on leaf lamina in concentric rings, **(G)** susceptible line Q38 with concentric ring-like sporulation patterns characteristic of *Pyrenopeziza brassicae*.

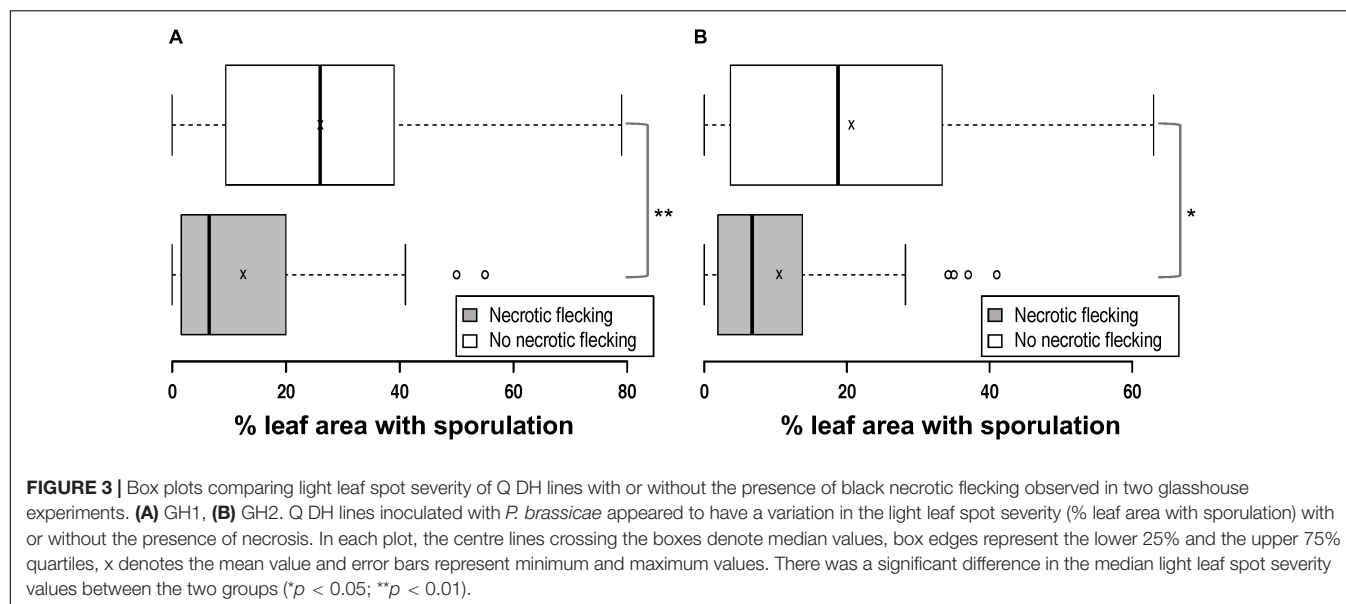
DH lines with or without a necrotic response was statistically significant in GH1 ($p < 0.01$) and GH2 ($p < 0.05$) experiments.

Mapping of Quantitative Trait Locus for Resistance Against *Pyrenopeziza brassicae*

Single marker analysis, which was used as the initial approach for the genetic mapping of the resistance against *P. brassicae*, indicated several loci with significant marker-trait associations. IM analysis within individual experiments identified four QTLs for resistance against *P. brassicae* across three linkage groups in the genetic linkage map for the Q DH population. Summary of the QTLs detected in each experiment, including maximum values of LOD scores, QTL positions, the percentage of phenotypic variance explained, and the estimate of QTL effects, is given in **Table 2**. Identification of QTL positions and effects were compared between transformed and untransformed data. This indicated similar results except for one QTL maximum on linkage group A10 for which the LOD score was less than the significant threshold with untransformed data. This QTL was removed from further analysis. One major QTL exceeding the LOD threshold was detected in the GH1 experiment located on the linkage group C6 (maximum LOD at 31.5 cM), accounting for 33.8% of the phenotypic variance. In comparison, three and two QTLs were detected in GH2 and CE experiments, respectively. The QTLs identified in the GH2 experiment were located on the linkage groups C1 (maximum LOD at 30.3 cM), C6 (maximum LOD at

31.5 cM), and C9 (maximum LOD at 57.0 cM), accounting for between 24.31 and 52.74% phenotypic variance. The two QTLs detected in the CE experiment included a major QTL located on the linkage group C1 (maximum LOD at 30.3 cM), accounting for 69.37% of the variance identified, and a QTL with a relatively small effect on the linkage group C6 (maximum LOD at 49.0 cM), accounting for 20.23% of the variance identified. Comparing the QTLs identified within different experiments (**Figure 4**), one of the QTLs on linkage group C6 was detected in both GH1 and GH2 experiments, and the QTL on linkage group C1 was detected in both GH2 and the CE experiments. For the GH1 experiment, a putative QTL co-located with the QTL on linkage group C1 identified in GH2 and CE experiments was detected with a LOD score of 3.1 that was just below the significance threshold ($\text{LOD} = 3.3$). QTL analysis for the combined data across all three experiments detected three QTLs co-located with those identified within individual experiments. These included QTLs located on the linkage groups C1 (maximum LOD at 30.4 cM), C6 (maximum LOD at 31.4 cM), and C9 (maximum LOD at 56.8 cM) (**Figure 4**), accounting for 37.31, 37.22, and 22.83% of phenotypic variance, respectively.

Second QTL analysis was done for GH1 and GH2 experiments by taking sub-populations of Q DH lines based on the presence or absence of necrosis with the intention of distinguishing QTL effects related to these two phenotype groups. According to the results, only the QTL identified on linkage group C1 remained significant in GH2 for the sub-population of Q DH lines without necrosis. The QTL on the linkage group C9 was also retained,



with a small decrease in the LOD score. The same phenomenon was observed in GH1, where the putative QTL maximum on the linkage group C1 was detected with the sub-population of Q DH lines without necrosis. Therefore, it is possible that the QTLs on the linkage groups C1 and C9 contribute to the reduced sporulation without necrosis. QTL mapping of the sub-population of DH lines with necrosis retained the QTL maximum on linkage group C6 with slightly reduced LOD scores for both GH1 and GH2, while showing a considerable loss of QTL maxima on the linkage group C1. The overall phenotypic variance in the sub-population of DH lines with necrosis may be attributed to the combined effects of QTLs (i.e., those related to reduced sporulation and necrosis). Considering the QTL maxima identified in the sub-population, it can be suggested that the QTL on the linkage group C6 contributes more toward the phenotypic variance in the group of DH lines with necrosis. However, no significant interactions with marker data exceeding the QTL threshold were detected for the binary data on necrosis.

Assessment of the Sub-Cuticular Growth Phase of *Pyrenopeziza brassicae* in Q Doubled Haploid Lines

Using the results obtained in the three phenotyping experiments (GH1, GH2, and CE), four Q DH lines were selected to represent differences in the amounts of sporulation and the necrosis observed. These included Q83 with an average of <1% leaf area affected with sporulation in the presence of necrosis, Q4 with <5% leaf area with sporulation in the presence of necrosis, Q69 with <2% leaf area with sporulation without necrosis, and Q38 with >59% of leaf area with sporulation without necrosis. Imola and Bristol were the resistant and susceptible control cultivars, respectively. Analysis of *P. brassicae* DNA content in leaf discs taken from the points of inoculation showed a significant increase between 0 and 24 dpi in all the lines and cultivars. There were significant differences between

lines/cultivars ($p < 0.05$). Position and parallelism regression analysis indicated three distinct groups based on the difference in the increase in *P. brassicae* DNA over time: group 1 contained Q38, group 2 consisted of Q04, Q83, Q69, and cv. Bristol, and group 3 contained cv. Imola (Figure 5). Line Q38, which developed large numbers of *P. brassicae* acervuli in all three phenotyping experiments (GH1, GH2, and CE), had the greatest amount of *P. brassicae* DNA and the greatest rate of increase over time. Cultivar Imola had significantly less *P. brassicae* DNA than the rest of the lines/cultivars and the amount of DNA increased over time at a slower rate.

Scanning electron micrographs also indicated that *P. brassicae* was capable of infecting and colonising all the lines selected in this experiment even though the extent of sub-cuticular hyphal growth varied among them (Figure 6). Lines Q83, Q4, and Q69 had hyphae growing more prominently along the leaf veins at early time points (i.e., 3–7 dpi). In addition, sub-cuticular hyphal growth on these lines appeared to follow the branching patterns of the main and lateral veins. This observation was consistent with the patterns of necrosis and acervuli production on the resistant lines. As the pathogen colonisation progressed with time (i.e., 14 dpi), the hyphae appeared to spread out onto the leaf lamina to a certain extent and more fungal biomass could be observed on the leaf lamina at c. 24 dpi. In addition, epidermal cell collapse associated with hyphae was observed on the Q DH lines with necrosis. In contrast, extensive hyphal growth could be observed on both the leaf veins and leaf lamina of the susceptible line Q38 from 7 dpi with more hyphae branching out from leaf veins onto the leaf lamina.

DISCUSSION

This article reports identification of new QTLs for resistance against *P. brassicae* derived from a *B. napus* secondary gene pool (ancestral genomes). Results from the phenotypic and genetic

TABLE 2 | Quantitative trait loci (QTLs) detected across three phenotyping experiments for resistance against *P. brassicae* in the *B. napus* Q DH population.

| Experiment ^a | Linkage group ^b | QTL position (cM) ^c | Support interval (cM) ^d | Peak LOD | R ² (%) ^e | Additive effect | QTL significance ^f |
|-------------------------|----------------------------|--------------------------------|------------------------------------|----------|---------------------------------|-----------------|-------------------------------|
| GH1 | C6 | 31.5 | 29.3–32.9 | 3.4 | 33.8 | 10.98 | *** |
| GH2 | C1 | 30.3 | 30.3–31.8 | 5.5 | 52.74 | 12.03 | **** |
| | C6 | 31.5 | 29.6–32.9 | 3.5 | 34.59 | 9.15 | *** |
| | C9 | 57.0 | 56.2–58.6 | 4.1 | 24.31 | 7.54 | **** |
| CE | C1 | 30.3 | 30.1–30.5 | 7.4 | 69.37 | 12.21 | ** |
| | C6 | 49.0 | 47.6–49.9 | 3.3 | 20.23 | –4.77 | *** |

^aDH population was phenotyped in two glasshouse experiments (GH1 and GH2) and a controlled-environment experiment (CE), and QTL was determined within each individual experiment. A summary of the phenotyping experiments is given in Table 1.

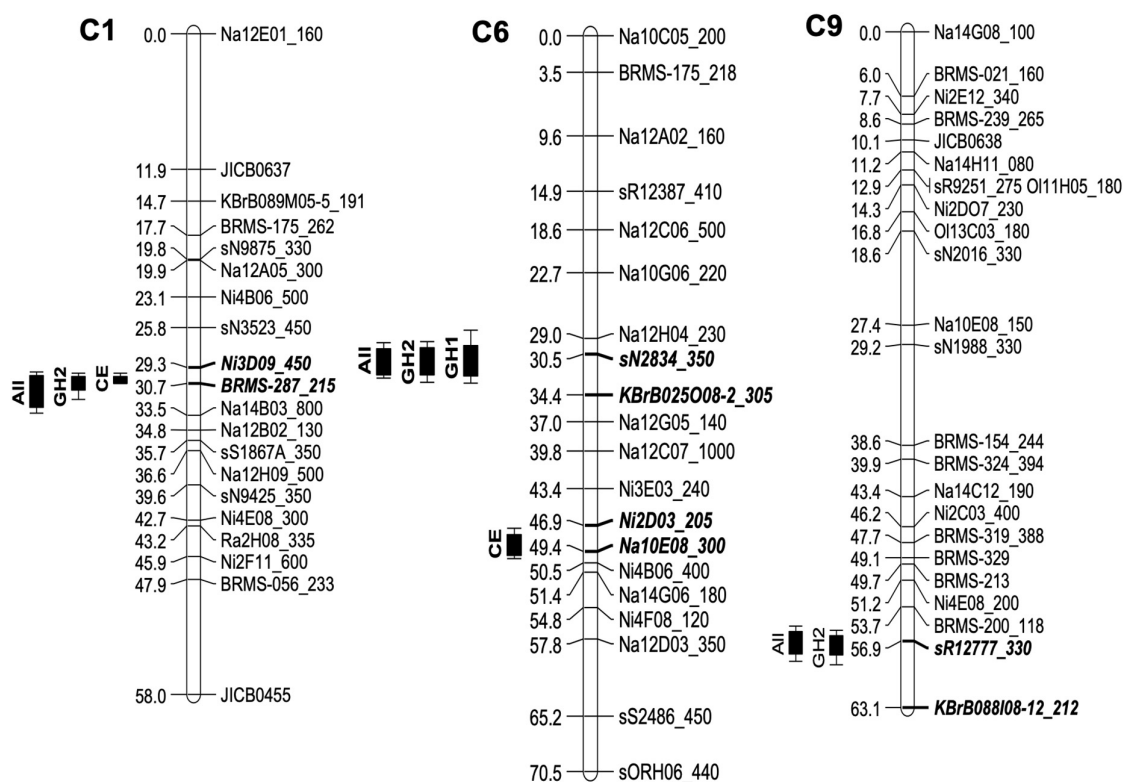
^bLinkage groups are labelled according to the standard chromosome/linkage group nomenclature of *B. napus* (A1–A10 and C1–C9) agreed by the Multinational Brassica Genome Project (MBGP) Steering Committee (<http://www.brassica.info/resource/maps/lg-assignments.php>).

^cFlanking markers for each QTL have been indicated in Figure 4.

^dLOD-1.5 support interval, which has a confidence interval close to 95%, is given for each of the QTL detected.

^e% phenotypic variance explained by the QTL.

^fSignificance based on individual marker-trait association. Significance at the 1, 0.1, and 0.01% levels are indicated by **, *** and ****, respectively.

**FIGURE 4** | Quantitative trait loci (QTLs) for resistance against *P. brassicae* in the *B. napus* Q DH population detected using three phenotyping experiments.

Phenotyping of the Q DH population was done using two glasshouse experiments (GH1 and GH2) and a CE experiment with each involving different *P. brassicae* populations. Four QTLs were identified across three linkage groups, C1, C6, and C9 using QTL analysis within individual experiments. Combined data across all three experiments identified three QTLs (labelled "All") co-locating with those identified within individual experiments. On the left, QTL positions are marked with LOD support intervals and flanking markers for each QTL are indicated by bold, italicised letters on the linkage maps.

analysis of host resistance against *P. brassicae* provided good evidence for the segregation of resistance against *P. brassicae* in the Q DH population. Q DH lines differed from each other in their ability to limit *P. brassicae* asexual sporulation. There were significant differences in the % leaf area covered with sporulation between different lines. Reduced *P. brassicae* sporulation appeared to be present with or without a necrotic

response and there was a significant difference between the two groups in the median values of % leaf area covered with acervuli. The concentric ring-like arrangement of the necrotic spots on the leaf lamina may indicate that *P. brassicae* asexual sporulation, which occurs in concentric rings in susceptible responses (Karandeni Dewage et al., 2018), is prevented by operation of the host resistance in these lines. This suggests

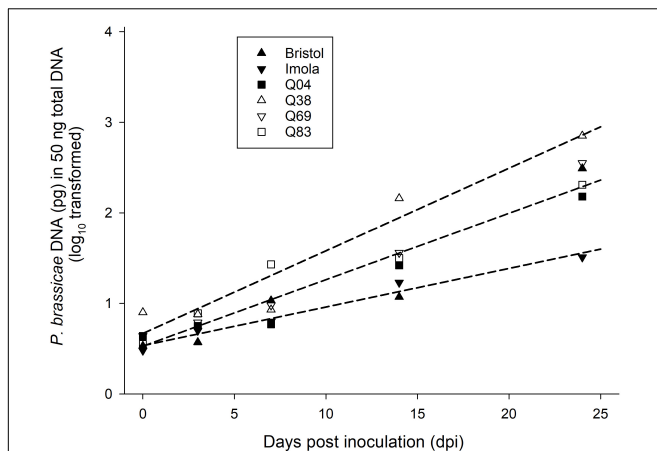


FIGURE 5 | Change with time in amount of *P. brassicae* DNA in selected *B. napus* Q DH lines and control cultivars. In a controlled-environment experiment, four Q DH lines differing in resistance and oilseed rape cultivars Bristol and Imola were point-inoculated with a *P. brassicae* conidial suspension. Amounts of *P. brassicae* DNA in leaf samples taken from the points of inoculation between 0 and 24 dpi were quantified using qPCR. Position and parallelism regression analysis was used to analyse the difference in the increase of DNA with time between the six lines/cultivars. Data were best fitted by three regression lines and there were three distinct groups based on the difference in the increase of DNA with time: group 1 contained Q38 ($y = 0.09x + 0.67$), group 2 consisted of Q04, Q83, Q69, and cv. Bristol ($y = 0.07x + 0.53$) and group 3 contained cv. Imola ($y = 0.04x + 0.53$).

that host recognition occurs at a late stage (c. 10–14 dpi) of *P. brassicae* colonisation, possibly during the asexual sporulation phase. However, the variation in amount of sporulation, observed between Q DH lines that gave a necrotic response upon pathogen recognition, suggests that there may be a background resistance or other resistance genes segregating in this population that affect the overall level of resistance to the pathogen.

Of the parental lines, neither the synthetic *B. napus* nor the C sub-genome parent of the synthetic *B. napus* (*Brassica oleracea altantica*) were available to test in our experiments. However, in the GH1 experiment, we were able to include *B. rapa olifera* “29” (A sub-genome parent of the synthetic *B. napus*) and cv. Tapidor, which showed moderate to high percentage leaf area with *P. brassicae* asexual sporulation without a necrotic response. According to these observations, it can be speculated that most of the favourable alleles for resistant QTLs were likely to have been contributed by C sub-genome parent via synthetic *B. napus*. The Q DH population is known to have an asymmetric distribution of marker polymorphisms between the A and C sub-genomes with allelic diversity diverted more toward the C sub-genome (Smooker et al., 2011). This seems to be true for the segregation of resistance against *P. brassicae* in the Q DH population, considering the distribution of QTLs on the C sub-genome.

In three phenotyping experiments, four QTLs were identified with moderate to large QTL effects. Linkage groups C1 and C6 appeared to have co-locating QTL stable across GH2/CE and GH1/GH2 experiments, respectively. Of these, the QTL on the linkage group C1 appeared to have a major effect on limiting *P. brassicae* asexual sporulation. QTL analysis is considered as

the initial step toward marker-assisted selection (MAS) in plant breeding, and for a particular QTL to be effective in a plant breeding programme, it is important to confirm the repeatability and the efficiency of the QTLs in different environments (Collard et al., 2005). Accurate identification of QTL depends on the quality of the phenotyping data and the robustness of the linkage map. A significant positive correlation of the phenotype data between the different experiments indicated consistency in disease scoring methodology across experiments. However, it should be noted that the relatively small population size used in this study probably resulted in false negatives, particularly in detecting QTLs with relatively small effects, along with possible over-estimation of the QTL effects.

There seem to be several QTLs contributing to the overall phenotypic variation. QTL-mediated resistance associated with reduced *P. brassicae* sporulation or leaf necrosis is probably controlled by different *B. napus* resistance loci. However, no significant QTL was detected when the binary data for necrosis were used as a phenotype, which could have been due to the genetic complexity of this phenotype. All the experiments were done with *P. brassicae* populations (mixtures of isolates) and some of the lines that had necrosis in GH1 did not produce necrosis in GH2 and *vice versa*. It is possible that there were different pathogen races or effectors recognised by the host. Furthermore, there seemed to be differences in the intensity of black flecking observed in different lines, indicating that the expression of this necrosis phenotype may be affected by other QTLs for resistance, or that this could be a component of a network of host responses. When we separate the DH lines based on the presence or absence of necrosis, the lines with necrosis contain phenotypic variation attributed to the combined effect of different loci, whereas the effect of necrosis is eliminated from the lines without necrosis. This may provide a possible genetic explanation for the significantly smaller median LLS severity values observed in the group of DH lines that showed necrosis compared to the group without necrosis.

Similar phenotype classes for resistance against *P. brassicae* have been explained by Bradburne et al. (1999) in a DH population of *B. napus*. Bradburne et al. (1999) reported two major genes for resistance against *P. brassicae*. Linkage analysis positioned the gene corresponding to “no asexual sporulation” (*PBR1*) on linkage group A1 and the gene responsible for “dark necrotic flecking” (*PBR2*) on linkage group C6. The QTL on linkage group C6 has been localised toward the centre of the linkage group in both Bradburne et al. (1999) and the present study. There is a possibility that both the studies refer to the same resistance locus. However, the linkage map published by Bradburne et al. (1999) contains limited information with only two restriction fragment length polymorphism (RFLP) markers on C6. There are no common markers between the two studies, which are necessary to create a consensus map between different populations to enable the comparison of QTLs across different studies. Therefore, it is difficult to draw a conclusion with the information currently available. Regarding the work reported by Pilet et al. (1998) on quantitative resistance against *P. brassicae*, QTLs were detected mostly on the A sub-genome (linkage groups A2, A6, A7, and A9) and one QTL was detected on the

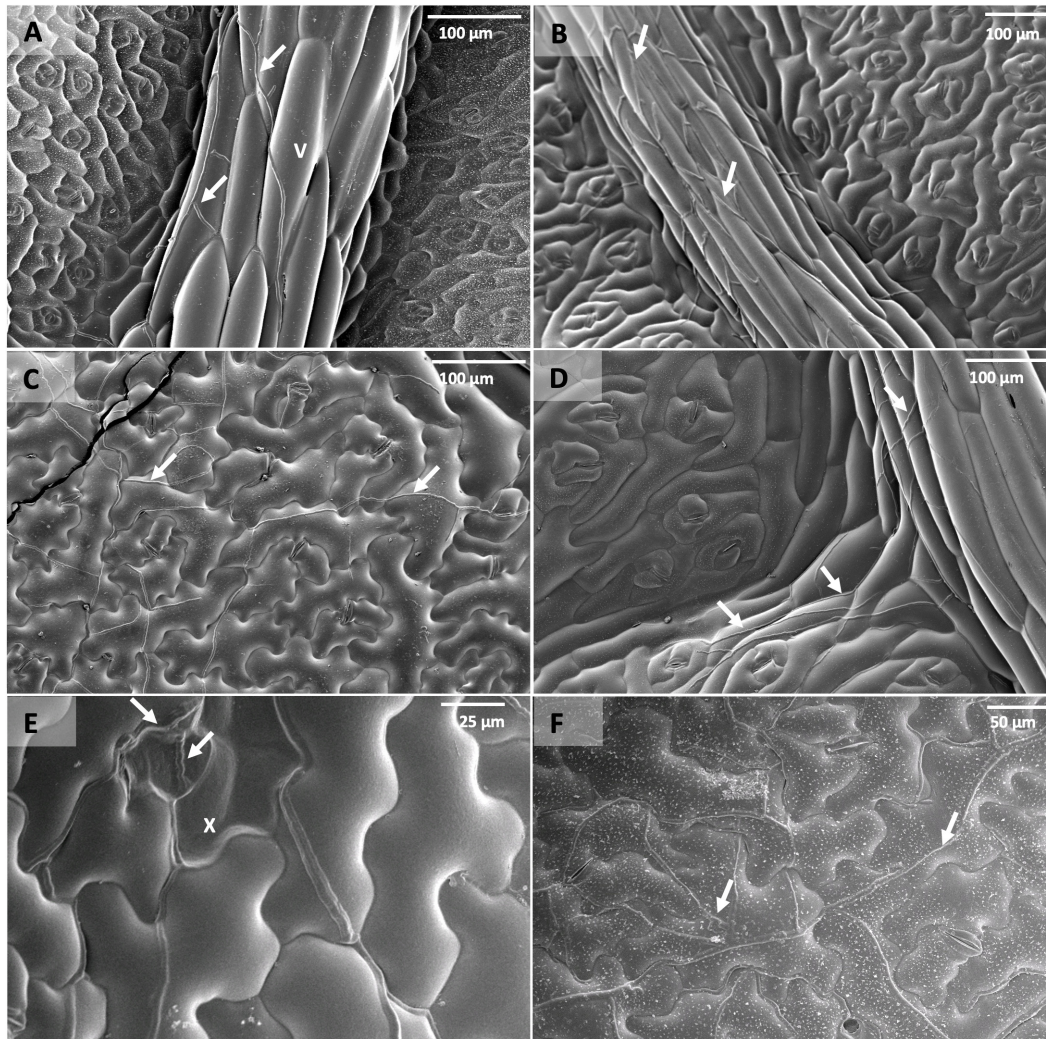


FIGURE 6 | Scanning electron micrographs of leaf discs taken from *B. napus* Q DH lines point-inoculated with *P. brassicae* conidia in a controlled environment experiment. **(A)** Q83 at 7 dpi, **(B)** Q4 at 7 dpi, **(C)** Q38 at 7 dpi, **(D)** Q04 at 7 dpi showing *P. brassicae* hyphae following leaf vein (v) branching patterns, **(E)** epidermal cell collapse (x) on Q83 associated with *P. brassicae* hyphae, **(F)** Q4 at 24 dpi. Arrows indicate *P. brassicae* hyphae growing in sub-cuticular spaces of leaf vein and leaf lamina tissues.

C sub-genome (linkage group C4), whereas the present study identified QTLs on linkage groups C1, C6, and C9.

Using a DH population derived from the material described by Bradburne et al. (1999) and Boys et al. (2012) reported a single locus for resistance corresponding to the black flecking phenotype (*PBR2*) that mapped to the bottom end of chrA1. The second major gene (*PBR1*, corresponding to the absence of asexual sporulation) reported by Bradburne et al. (1999) was not identified and it has been suggested that *PBR1* might have been lost during the breeding process. In the present study, we have identified DH lines with little to no sporulation without necrosis. These lines can be used to further dissect the genetic basis of this resistant phenotype by developing a larger mapping population for fine-scale mapping. Availability of *Brassica napus* genomic resources offers new possibilities for the identification of host resistance genes and provides molecular tools to assist

in marker-assisted selection (MAS) for disease resistance. There are a few *B. napus* genome sequences published, including the genome sequence of cv. Tapidor (Chalhoub et al., 2014; Bayer et al., 2017; Sun et al., 2017), which is one of the parental lines of the Q DH population. Sequencing of the flanking markers can be used to identify the corresponding genomic regions of the QTL on the *B. napus* genome, facilitating the identification of candidate resistance genes.

The Q DH population segregates for vernalisation and winter hardiness (Smooker et al., 2011), making it difficult to assess some of the lines directly in winter oilseed rape field experiments in the United Kingdom. Therefore, we chose to phenotype the Q DH population under controlled-environment and glasshouse conditions to enable the identification of different components of resistance without being affected by other characters segregating in this population. Instead of using single-spore isolates, plants

were inoculated with a different *P. brassicae* population in each experiment, representing natural inoculum. In a separate study that assessed resistance against *P. brassicae* in different *B. napus* genotypes, selected Q DH lines have shown more resistance compared to that of commercial oilseed rape cultivars. In addition, the resistance in those lines appeared to be less sensitive toward the increasing virulence of *P. brassicae* isolates (Karandeni Dewage et al., 2021). This agrees with the stability of the resistance in Q DH lines across different *P. brassicae* populations observed in the present study. Host resistance QTLs that are stable across different *P. brassicae* populations are of particular importance to oilseed rape breeders.

Our results suggest that the resistance segregating in this DH population limits colonisation/sporulation by the pathogen rather than eliminating the pathogen. *Pyrenopeziza brassicae* was able to infect and colonise both resistant and susceptible Q DH lines and control cultivars with a significantly smaller rate of increase of pathogen biomass in resistant lines than that in susceptible lines. According to the qPCR data, three resistant Q DH lines were in the same group as cv. Bristol that had a significantly greater number of *P. brassicae* acervuli in phenotyping experiments. This also supports the suggestion that there may be resistance operating during the time of *P. brassicae* asexual sporulation. According to SEM, hyphal growth was more prominent along the leaf veins of resistant lines, especially at early time points. According to Boys et al. (2012), the amount of *P. brassicae* DNA was significantly greater in leaf vein/midrib tissues than elsewhere in the leaves of cv. Imola when samples were taken from the points of inoculation. This suggests that the pathogen was able to colonise more of leaf vein/midrib tissues than tissues of interveinal regions. It can be assumed that abundant vascular bundles in leaf veins provide the pathogen with more access to resources and hence support the extracellular colonisation of leaf veins, midribs, and petioles. This may also explain the presence of the black flecking phenotype, mainly along the leaf veins, midrib, or petioles, and the production of occasional acervuli along with these tissues in the case of resistant cultivars. There have been similar reports of other endophytic fungi with more affinity toward leaf vein/petiole colonisation (Photita et al., 2001; Toofanee and Dulymamode, 2002).

Even though LLS is considered as a major disease problem of oilseed rape in the UK with many epidemics since 1970s, little is known about this pathosystem in contrast to that of other important diseases, such as phoma stem canker (Boys et al., 2007; Karandeni Dewage et al., 2018). So far, there have been only two published studies on genetic mapping of major-gene-mediated resistance (Bradburne et al., 1999; Boys et al., 2012) and one study reporting quantitative resistance against *P. brassicae* in *B. napus* (Pilet et al., 1998). Therefore, new studies on qualitative and quantitative resistance against *P. brassicae* are invaluable to combat LLS to sustain oilseed rape production in the United Kingdom.

Quantitative trait locus mapping can be used to determine QTL effects, interactions between resistance genes, race-specificity of resistance, etc., providing insights into resistance against pathogens that has complex inheritance (Young, 1996; St. Clair, 2010). Even though QDR has been studied in several

pathosystems, underlying molecular mechanisms of QDR are not well understood, with few studies reporting functional characterisation of resistance QTL. In this study, we were able to do detailed phenotyping of the resistance against *P. brassicae* originating from the ancestral brassica species and allocate the observed variation into its genetic components. Resistant lines identified can serve as pre-breeding material and the QTLs that are stable across different experiments could be utilised in MAS in oilseed rape breeding programmes to improve cultivar resistance against *P. brassicae* with further resolution of resistance QTLs. Associated markers can be used as a starting point for functional identification of the genes controlling resistance against *P. brassicae* that can contribute to our knowledge on mechanisms of partial resistance/QDR of crops against pathogens.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

CKD, HS, Y-JH, and BF conceived the project. CKD and KC designed and did the experiments. AQ and CKD did the data analysis. RW contributed to the resources and assisted with analysis. CKD conceptualised and drafted the manuscript with additions and edits from BF, AQ, Y-JH, and HS. All authors read and approved the final manuscript.

FUNDING

The authors acknowledge funding support from AHDB Cereals and Oilseeds (grant/award number: 2140010), UK Biotechnology and Biological Sciences Research Council (grant/award numbers: BBSRC/BB/P00489X/1 and BB/V01725X/1-BBSRC Follow-on Fund), Innovate UK (grant/award number: 102641), Department for Environment, Food and Rural Affairs (OREGIN) (grant/award number: CH0110), Gen Foundation and the University of Hertfordshire.

ACKNOWLEDGMENTS

We thank Mark Nightingale (Elsoms Seeds UK Ltd.) and Vasilis Gegas (Limagrain UK Ltd.) for providing seeds of Q DH lines for this study. We also thank Kavithra Wijerathna and Fraser Martin for their help with experimental work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.786189/full#supplementary-material>

REFERENCES

- AHDB (2021). *Light Leaf Spot*. Available online at: <https://ahdb.org.uk/lightleafspot> (accessed January 02, 2021).
- Bayer, P. E., Hurgobin, B., Golicz, A. A., Chan, C. K. K., Yuan, Y., Lee, H., et al. (2017). Assembly and comparison of two closely related *Brassica napus* genomes. *Plant Biotechnol. J.* 15, 1602–1610. doi: 10.1111/pbi.12742
- Boyd, L. A., Ridout, C., O'Sullivan, D. M., Leach, J. E., and Leung, H. (2013). Plant–pathogen interactions: disease resistance in modern agriculture. *Trends Genet.* 29, 233–240. doi: 10.1016/j.tig.2012.10.011
- Boys, E. F., Roques, S. E., Ashby, A. M., Evans, N., Latunde-Dada, A. O., Thomas, J. E., et al. (2007). Resistance to infection by stealth: *Brassica napus* (winter oilseed rape) and *Pyrenopeziza brassicae* (light leaf spot). *Eur. J. Plant Pathol.* 118, 307–321. doi: 10.1007/s10658-007-9141-9
- Boys, E. F., Roques, S. E., West, J. S., Werner, C. P., King, G. J., Dyer, P. S., et al. (2012). Effects of R gene-mediated resistance in *Brassica napus* (oilseed rape) on asexual and sexual sporulation of *Pyrenopeziza brassicae* (light leaf spot). *Plant Pathol.* 61, 543–554. doi: 10.1111/j.1365-3059.2011.02529.x
- Bradburne, R., Majer, D., Magreth, R., Werner, C., Lewis, B., and Mithen, R. (1999). Winter oilseed rape with high levels of resistance to *Pyrenopeziza brassicae* derived from wild *Brassica* species. *Plant Pathol.* 48, 550–558. doi: 10.1046/j.1365-3059.1999.00373.x
- Brown, J. (2004). *Experimental Design Generator and Randomiser II*. Available online at: <http://www.edgarweb.org.uk> (accessed May 06, 2017).
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435
- Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142, 169–196. doi: 10.1007/s10681-005-1681-5
- CropMonitor (2016). *Survey of Commercially Grown Winter Oilseed Rape*. Department for Environment, Food and Rural Affairs. Available online at: www.cropmonitor.co.uk/wosr/surveys/wosr.cfm (accessed August 1, 2016).
- DEFRA (2020). *Farming Statistics – Final Crop Areas, Yields, Livestock Populations and Agricultural Workforce at 1 June 2020 United Kingdom*. Available online at: <https://www.gov.uk/government/statistics> (accessed November 10, 2021).
- Doebley, J. F., Gaut, B. S., and Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell* 127, 1309–1321. doi: 10.1016/j.cell.2006.12.006
- Evans, N., Baierl, A., Brain, P., Welham, S. J., and Fitt, B. D. L. (2003). Spatial aspects of light leaf spot (*Pyrenopeziza brassicae*) epidemic development on winter oilseed rape (*Brassica napus*) in the United Kingdom. *Phytopathology* 93, 657–665. doi: 10.1094/PHYTO.2003.93.6.657
- French, E., Kim, B. S., and Iyer-Pascuzzi, A. S. (2016). Mechanisms of quantitative disease resistance in plants. *Semin. Cell Dev. Biol.* 56, 201–208. doi: 10.1016/j.semcdb.2016.05.015
- Haley, C. S., and Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69, 315–324. doi: 10.1038/hdy.1992.131
- Hyten, D. L., Song, Q., Zhu, Y., Choi, I. Y., Nelson, R. L., Costa, J. M., et al. (2006). Impacts of genetic bottlenecks on soybean genome diversity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 16666–16671. doi: 10.1073/pnas.0604379103
- Karandeni Dewage, C. S., Klöppel, C. A., Stotz, H. U., and Fitt, B. D. L. (2018). Host–pathogen interactions in relation to management of light leaf spot disease (caused by *Pyrenopeziza brassicae*) on *Brassica* species. *Crop Pasture Sci.* 69, 9–19. doi: 10.1071/CP16445
- Karandeni Dewage, C. S., Qi, A., Stotz, H. U., Huang, Y. J., and Fitt, B. D. L. (2021). Interactions in the *Brassica napus*–*Pyrenopeziza brassicae* pathosystem and sources of resistance to *P. brassicae* (light leaf spot). *Plant Pathol.* 70, 2104–2114. doi: 10.1111/ppa.13455
- Karolewski, Z., Fitt, B. D. L., Latunde-Dada, A. O., Foster, S. J., Todd, A. D., Downes, K., et al. (2006). Visual and PCR assessment of light leaf spot (*Pyrenopeziza brassicae*) on winter oilseed rape (*Brassica napus*) cultivars. *Plant Pathol.* 55, 387–400. doi: 10.1111/j.1365-3059.2006.01383.x
- Katche, E., Quezada-Martinez, D., Katche, E. I., Vasquez-Teuber, P., and Mason, A. S. (2019). Interspecific hybridization for *Brassica* crop improvement. *Crop Breed. Genet. Genomics* 1:e190007. doi: 10.20900/cbgg20190007
- Maddock, S. E., Ingram, D. S., and Gilligan, C. A. (1981). Resistance of cultivated brassicas to *Pyrenopeziza brassicae*. *Trans. Br. Mycol. Soc.* 76, 371–382. doi: 10.1016/S0007-1536(81)80063-0
- Mithen, R. F., and Magrath, R. (1992). Glucosinolates and resistance to *Leptosphaeria maculans* in wild and cultivated *Brassica* species. *Plant Breed.* 108, 60–68. doi: 10.1111/j.1439-0523.1992.tb00100.x
- Neik, T. X., Barbetti, M. J., and Batley, J. (2017). Current status and challenges in identifying disease resistance genes in *Brassica napus*. *Front. Plant Sci.* 8:1788. doi: 10.3389/fpls.2017.01788
- Parsad, R., Gupta, V. K., Batra, P. K., Satpati, S. K., and Biswas, P. (2007). *Monograph on Alpha Designs*. New Delhi: ICAR-IASRI.
- Payne, R. W., Harding, S. A., Murray, D. A., Soutar, D. M., Baird, D. B., Glaser, A. I., et al. (2011). *The Guide to Genstat Release 14, Part 2: Statistics*. Hemel Hempstead: VSN International.
- Photita, W., Lumyong, S., Lumyong, P., and Hyde, K. D. (2001). Endophytic fungi of wild banana (*Musa acuminata*) at Doi Suthep Pui National Park, Thailand. *Mycol. Res.* 105, 1508–1513. doi: 10.1017/S0953756201004968
- Pilet, M. L., Delourme, R., Foisset, N., and Renard, M. (1998). Identification of QTL involved in field resistance to light leaf spot (*Pyrenopeziza brassicae*) and blackleg resistance (*Leptosphaeria maculans*) in winter rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 97, 398–406. doi: 10.1007/s001220050909
- Poland, J. A., Balint-Kurti, P. J., Wisser, R. J., Pratt, R. C., and Nelson, R. J. (2009). Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.* 14, 21–29. doi: 10.1016/j.tplants.2008.10.006
- Rawlinson, C. J., Sutton, B. C., and Muthyalu, G. (1978). Taxonomy and biology of *Pyrenopeziza brassicae* sp. nov. (*Cylindrosporium concentricum*), a pathogen of winter oilseed rape (*Brassica napus* ssp. *oleifera*). *Trans. Br. Mycol. Soc.* 71, 425–439. doi: 10.1016/S0007-1536(78)80070-9
- Robin, A. H. K., Larkan, N. J., Laila, R., Park, J. I., Ahmed, N. U., Borhan, H., et al. (2017). Korean *Brassica oleracea* germplasm offers a novel source of qualitative resistance to blackleg disease. *Eur. J. Plant Pathol.* 149, 611–623. doi: 10.1007/s10658-017-1210-0
- Roux, F., Voisin, D., Badet, T., Balagué, C., Barlet, X., Huard-Chauveau, C., et al. (2014). Resistance to phytopathogens e tutti quanti: placing plant quantitative disease resistance on the map. *Mol. Plant Pathol.* 15, 427–432. doi: 10.1111/mpp.12138
- Silva, L. D. C. E., Wang, S., and Zeng, Z. B. (2012). “Composite interval mapping and multiple interval mapping: procedures and guidelines for using windows QTL Cartographer,” in *Quantitative Trait Loci (QTL): Methods and Protocols Methods in Molecular Biology*, ed. S. A. Rifkin (New York, NY: Humana Press), 75–119.
- Smooker, A. M., Wells, R., Morgan, C., Beaudoin, F., Cho, K., Fraser, F., et al. (2011). The identification and mapping of candidate genes and QTL involved in the fatty acid desaturation pathway in *Brassica napus*. *Theor. Appl. Genet.* 122, 1075–1090. doi: 10.1007/s00122-010-1512-5
- Snowdon, R. J., and Iniguez Luy, F. L. (2012). Potential to improve oilseed rape and canola breeding in the genomics era. *Plant Breed.* 131, 351–360. doi: 10.1111/j.1439-0523.2012.01976.x

- St. Clair, D. A. (2010). Quantitative disease resistance and quantitative resistance loci in breeding. *Annu. Rev. Phytopathol.* 48, 247–268.
- Sun, F., Fan, G., Hu, Q., Zhou, Y., Guan, M., Tong, C., et al. (2017). The high-quality genome of *Brassica napus* cultivar ‘ZS 11’ reveals the introgression history in semi-winter morphotype. *Plant J.* 92, 452–468. doi: 10.1111/tpj.13669
- Sylvester-Bradley, R., Makepeace, R. J., and Broad, H. (1984). A revised code for stages of development in oilseed rape (*Brassica napus* L.). *Aspects Appl. Biol.* 6, 399–419.
- Toofanee, S. B., and Dulyamode, R. (2002). Fungal endophytes associated with *Cordemoya integrifolia*. *Fungal Divers.* 11, 169–175.
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78.
- Wang, S., Basten, C. J., and Zeng, Z. B. (2012). *Windows QTL Cartographer 2.5*. Raleigh, NC: Department of Statistics, North Carolina State University.
- Young, N. D. (1996). QTL mapping and quantitative disease resistance in plants. *Annu. Rev. Phytopathol.* 34, 479–501. doi: 10.1146/annurev.phyto.34.1.479

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 Karandeni Dewage, Cools, Stotz, Qi, Huang, Wells and Fitt. 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.



Single Nucleotide Polymorphism Detection for Peach Gummosis Disease Resistance by Genome-Wide Association Study

OPEN ACCESS

Edited by:

Jianjun Chen,
University of Florida, United States

Reviewed by:

Ainong Shi,
University of Arkansas, United States
Jian Guo,
Shandong Agricultural University,
China
Alice Kujur,
International Crops Research Institute
for the Semi-Arid Tropics (ICRISAT),
India

*Correspondence:

Zhengwen Ye
yezhenqwen1300@163.com
Zhongshan Gao
gaozhongshan@zju.edu.cn

[†] 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: 24 August 2021

Accepted: 28 December 2021

Published: 07 February 2022

Citation:

Li X, Wang J, Su M, Zhou J,
Zhang M, Du J, Zhou H, Gan K, Jin J,
Zhang X, Cao K, Fang W, Wang L,
Jia H, Gao Z and Ye Z (2022) Single
Nucleotide Polymorphism Detection
for Peach Gummosis Disease
Resistance by Genome-Wide
Association Study.
Front. Plant Sci. 12:763618.
doi: 10.3389/fpls.2021.763618

Xiongwei Li^{††}, Jiabo Wang^{2†}, Mingshen Su¹, Jingyi Zhou³, Minghao Zhang¹, Jihong Du¹,
Huijuan Zhou¹, Kexin Gan⁴, Jing Jin⁴, Xianan Zhang¹, Ke Cao⁵, Weichao Fang⁵,
Lirong Wang⁵, Huijuan Jia⁴, Zhongshan Gao^{4*} and Zhengwen Ye^{1*}

¹ Forest and Fruit Tree Institute, Shanghai Academy of Agricultural Sciences, Shanghai, China, ² Key Laboratory of Qinghai-Tibetan Plateau Animal Genetic Resource Reservation and Utilization (Southwest Minzu University), Ministry of Education, Chengdu, China, ³ Horticultural Department, Shanghai Municipal Agricultural Technology Extension and Service Center, Shanghai, China, ⁴ Key Laboratory for Horticultural Plant Growth, Department of Horticulture, Development and Quality Improvement of State Agriculture Ministry, Zhejiang University, Hangzhou, China, ⁵ Zhengzhou Fruit Research Institute, Chinese Academy of Agriculture Sciences, Zhengzhou, China

Peach gummosis is one of the most widespread and destructive diseases. It causes growth stunting, yield loss, branch, trunk, and tree death, and is becoming a restrictive factor in healthy and sustainable development of peach production. Although a locus has been identified based on bi-parental quantitative trait locus (QTL) mapping, selection of gummosis-resistant cultivars remains challenging due to the lack of resistant parents and of the complexity of an inducing factor. In this study, an integrated approach of genome-wide association study (GWAS) and comparative transcriptome was used to elucidate the genetic architecture associated with the disease using 195 accessions and 145,456 genome-wide single nucleotide polymorphisms (SNPs). The broad-sense and narrow-sense heritabilities were estimated using 2-year phenotypic data and genotypic data, which gave high values of 70 and 73%, respectively. Evaluation of population structure by neighbor-joining and principal components analysis (PCA) clustered all accessions into three major groups and six subgroups, mainly according to fruit shape, hairy vs. glabrous fruit skin, pedigree, geographic origin, and domestication history. Five SNPs were found to be significantly associated with gummosis disease resistance, of which SNPs285957, located on chromosome6 across 28 Mb, was detected by both the BLINK and the FarmCPU model. Six candidate genes flanked by or harboring the significant SNPs, previously implicated in biotic stress tolerance, were significantly associated with this resistance. Two highly resistant accessions were identified with low disease severity, which could be potential sources of resistance genes for breeding. Our results provide a fresh insight into the genetic control of peach gummosis disease.

Keywords: peach, gummosis disease, QTLs, genome-wide association study, candidate genes

INTRODUCTION

Peach [*Prunus persica* (L.) Batsch] is one of the most economically important deciduous fruit from the Rosaceae family (Li et al., 2013). It originated in northwest China, and has spread throughout China and the rest of the world because of its greater adaptability (Faust and Timon, 1995; Yu et al., 2018). But, the short-life syndrome due to gummosis is a long-lasting problem in the warm and moist climate regions.

Gummosis is a nonspecific disease response to pathogen infection, mechanical injury, drought and cold stress, or insect attack. It is characterized by a gum exudation from tree trunks, branches, and fruits in several fruit species, such as peach (Britton and Hendrix, 1982), almond (Popović et al., 2021), apricot (Liu et al., 2015), sweet cherry (Zhang L. et al., 2019), and in citrus (Fan et al., 2011). Gummosis in peach was first reported in central Georgia in 1974 (Weaver, 1974). The gum exudation on trunks, scaffold limbs, and branches significantly suppresses tree growth and fruit yield of susceptible peach varieties. It is one of the most destructive peach diseases in the south of China (Fan et al., 2011) and the southeastern United States (Weaver, 1974; Britton and Hendrix, 1982). Based on the conidial morphology, cultural characteristics, and nucleotide sequences, three *Botryosphaeria* fungi species were reported to be the main pathogens causing the peach gummosis disease: *Botryosphaeria dothidea* (anamorph *Fusicoccumaesculi*), *Botryosphaeria rhodina* (anamorph *Lasiodiplodia theobromae*), and *Botryosphaeria obtuse* (anamorph *Diplodiaseriata*) (Weaver, 1974). Of these, *Botryosphaeria dothidea* is the most common cause of the disease in a large number of hosts worldwide (Britton and Hendrix, 1982; Mancero-Castillo et al., 2018), while *Lasiodiplodia theobromae* has proven to be the most virulent, causing the largest lesions and most copious gummosis in China (Fan et al., 2011).

Previous studies on controlling peach gummosis disease have mainly involved chemical and biological controls with very limited efficacy. Therefore, the use and the breeding of gummosis resistance cultivars are the most cost-effective, environment-friendly, and healthy approach for long-term management of the disease (Beckman et al., 2011). As previously reported, although most peaches and nectarines are susceptible to gummosis disease to some degree, highly resistant genotypes also exist (Beckman and Reilly, 2005). However, using these resistant genotypes in breeding programs *via* conventional breeding methods remains a challenge due to the large plant size, self-compatibility, low genetic diversity, and the most restrictive factors, including the long juvenile periods and breeding cycles (Li et al., 2013; Aranzana et al., 2019). In addition, phenotypic variation of peach gummosis is always affected by several factors, such as wounding, pathogen infection, temperature, or humidity. As far as the genetic factor is concerned, a dominant allele for peach fungus gummosis resistance has been found in almond based on F₁ and BC₁F₁ population. Segregation and mapping analyses located the peach fungal gummosis resistance locus on chimeric linkage groups 6–8 near the leaf color locus (Mancero-Castillo et al., 2018). Furthermore, being the center of origin of peach, China has a huge population of wild relatives and landraces with high genetic diversity (Li et al., 2013; Micheletti et al.,

2015). These genetic resources always exhibit specific phenotypes of resistance and fruit quality but are rarely used in modern peach breeding programs. For example, *P. davidiana* carries resistance genes against the peach green aphid and can be used for aphid-resistant breeding (Li et al., 2019). It will be, therefore, helpful to perform extensive work with large-scale germplasm to elucidate the genetic mechanism controlling the severity of peach gummosis disease.

Next-generation sequencing technologies have not only promoted the development of genetics and genomics tools, but also greatly improved the understanding of the genetic basis of important agronomic traits. Based on RNA-seq technology, a large number of differentially expressed genes have been identified, which has enabled the elucidation of the molecular mechanism of plant-pathogen interaction of peach fungal gummosis after the infection (Gao et al., 2016). These genes have been found to be mainly involved in the process of cellular defense and metabolism of carbohydrates, the phenylpropanoid biosynthesis and metabolism pathway, anthocyanin biosynthetic pathway, and the ethylene and jasmonic biosynthetic pathways (Gao et al., 2016). Recently, the reactive oxygen species (ROS) production-scavenging system has been reported to play a crucial role in plant-pathogen interaction and in the development of gummosis caused by *Lasiodiplodia theobromae* (Zhang et al., 2020).

Genotyping by sequencing (GBS) is a method that combines the enzyme-based complexity reduction and the second-generation sequencing technology for marker discovery, with and without the reference genomes (Elshire et al., 2011). Despite the high rate of missing values in the GBS data, the advantages of simultaneous discovery of abundant single nucleotide polymorphisms (SNPs) at low cost, reduced the ascertainment bias compared with array-based markers, and a relatively easy automation still make it an efficient approach to detect polymorphisms and to identify various loci controlling traits both by biparental quantitative trait locus (QTL) mapping and by genome-wide association study (GWAS) (Elshire et al., 2011; Poland and Rife, 2012; Jarquín et al., 2014; Arruda et al., 2016; Minamikawa et al., 2018). A large number of studies on GWAS that were integrated with GBS have been reported in multiple plant species (Arruda et al., 2016; Cao et al., 2016, 2019; Guo et al., 2019; Siddique et al., 2019). By combining GBS-based QTL mapping with GWAS, 117 significant SNPs across the genome were identified to be associated with *P. capsici* root rot resistance in pepper (Siddique et al., 2019). Similarly, the genetic determinants of grape berry-related traits, including grape skin color, berry development period, berry weight, berry flesh texture, and berry flavor, were identified by performing GWAS with 179 grape accessions and 32,311 SNP markers derived from GBS analysis (Guo et al., 2019). Another example of GBS-based GWAS is where the candidate genes of 12 agronomic traits and selected domestication traits, including fruit shape, fruit color, fruit hairy, fruit weight, sorbitol, and catechin content, have been identified (Cao et al., 2016, 2019). Thus, keeping the above in view, an integrated approach of GWAS and comparative transcriptome was used in the present study. Here, the gummosis disease was scored in the large-scale peach core germplasm accessions, grown in the experimental field over the period of 2 years. The

plant resources were selected from the previous genetic diversity study (Li et al., 2013). A group of highly resistant accessions, especially the traditional landraces, were identified. These are potentially resistant parents to enrich the gene pool in modern peach breeding programs. The GWAS, combined with RNA-seq, was used to identify the associated SNP markers and candidate genes. The aim was to gain insights into the genetic basis of this complex trait and to apply the results in a peach genomic selection breeding program.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

A set of 195 peach accessions originating from 19 provinces and autonomous regions in China and United States, Italy, New Zealand, and Japan was selected (**Supplementary Table 1**). All trees were grafted on “MaoTao” rootstock and planted in the peach experimental trial fields of Shanghai Academy of Agricultural Sciences, Shanghai (N30°55′3.18″-E121°27′14.44″) during the March month of the year 2016. This region in China is characterized by high temperature and high humidity as the annual average temperature and humidity reach up to 17°C and 80%, respectively. The tree plants were managed under uniform conditions of irrigation, fertilization, and pest and disease control. Two accessions “Nan Shan Tian Tao” and “Sunfre” were additionally grown in two different locations as replications for resistance validation.

Evaluation of Lesions and Statistical Analysis of Gummosis Disease Score

The severity of peach gummosis was investigated in the end of the years 2018 and 2019. The score (0, 1, 3, 5, 7, and 9) for each tree was based on the number and area of gumming lesions on the trunks and limbs, a standard evaluation criterion adopted by the modern Chinese peach industry technology community. The minimum score of 0 refers to no visible symptoms or lesions on the whole tree, and the maximum 9 indicates a very severe infection on limbs and the main trunk. The scoring was as follows: 1, only 1–2 lesions with a diameter of the spots less than 3 cm identified on trunks or main limbs; 3, the lesion was from 1 or 2 spots covering an area of up to 25% of the whole plant with the gum spots not clearly distinguishable; 5, the total lesion area was 25–50% of the whole plant with the gum spots not clearly distinguishable; 7, the total lesion area covered from 50 to 75% of the whole plant with the gum spots not clearly distinguishable; and 9 when the total lesion area was more than 75% of the whole plant (**Supplementary Figure 1**). Those with mean scores of both years as stable to 1 or less were designated as high resistant, and those scores lower than or equal to 3 as middle resistant, and those greater than 3 were designated as susceptible.

The severity of disease was also compared in different peach groups (**Supplementary Table 1**) divided according to geographic origin and four phenotypes, including fruit pubescence (peach/nectarine), fruit shape (round/flat), fruit flesh color (red/yellow/white), and blossoming time (very early/early/middle/late). The statistical analyses, including

means, standard errors (SE), and the minimum and maximum values, were calculated using Graphpad Prism 8 software (Graphpad Software Inc., San Diego, California). Pearson correlation coefficients between the severity score of gummosis disease and geographic origin and four phenotypes were analyzed with the same software. The statistical significance was set at the $p < 0.001$ level. Ordinary one-way ANOVA and unpaired t -test (for fruit shape and hairy fruit skin) were used for paired and multiple comparisons, respectively.

Estimation of Best Linear Unbiased Prediction Values

Best linear unbiased prediction (BLUP) values were extracted from the 2-year (2018–2019) phenotypic data for gummosis disease using the linear mixed model in R-package lme4 based on the following equation:

$$Y_{ij} = \mu + A_i + y_j + e$$

where Y_{ij} is the vector of severity observation for each accession in each year, μ is the overall mean values for all individuals, A_i is the random effect of the i th individual accession ($i = 1, \dots, 195$), y_j is the random effect of the j th year ($j = 2018$ and 2019), and e is the residual error. Extraction of the random effects (accessions) in the model used the “ranef” function. The estimated BLUP values were used as phenotype values in the GWAS. In the equation, the ratio of the individuals’ (accessions’) variance in the total variance was used as estimated heritability (general heritability). The total variance was the sum variance of accessions, years, and e , that is the same as observations variance.

DNA Extraction, Re-sequencing, and Single Nucleotide Polymorphism Discovery

Young leaves with no disease from each accession were collected and frozen at -80°C . Total genomic DNA was isolated from 0.1-g tissue using the DNeasy 96 Plant Mini Kit (Qiagen, CA, United States) following the manufacturer’s protocol. The libraries with an insert size of 500 bp were constructed and sequenced by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China) using an Illumina HiSeq X Ten platform (Illumina, San Diego, CA) based on a paired-end mode, which resulted in sequenced fragments of 150 bp read length. The sequencing depth of each accession was greater than 10.33-fold with an average genome coverage of 98.14%. The raw sequencing data and SNP calling were analyzed using SAMTOOLS software (Li et al., 2009). The SNPs were filtered under the quality control parameter to remove those with more than a 7% individual missing rate and a minor allele frequency (MAF) that is lower than 0.05, according to the user manual of Beagle software 3.3.2.

Estimation of Population Structure, Genetics Parameters, and Genome-Wide Linkage Disequilibrium

The narrow sense of heritability of gummosis disease was estimated by GAPIT3 based on a mixed linear model using

whole of the marker data. Principal Component Analysis (PCA) and Neighbor Joining (NJ)-tree analysis were performed to find the clustered group and the genetic distance using GAPIT3 software (Wang and Zhang, 2021) for understanding the population structure. Eigen values and matrices were extracted as dimensionality reduction vectors from all genotype information. The first two PCs with major genetic variance were used to indicate population stratification. The clustered kinship was used to plot the NJ tree. To estimate the rate of linkage disequilibrium decay, r^2 values between each loci genotype were calculated using PopLD decay, which is a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files (Zhang C. et al., 2019). A window size, with averaged 300 kb across the whole genome, was used to calculate average r^2 values.

The genetic diversity indices for different populations, including observed heterozygosity (Ho), inbreeding coefficient (Fis), nucleotide diversity (π), and haplotypes, were calculated using the POPULATION program in the stacks package with a custom Perl script. Paired F-statistics values (*Fst*) (Weir and Cockerham, 1984) were calculated to measure the difference between populations using the same aforementioned program. Analysis of molecular variance (AMOVA) was used to partition the genetic variation into inter- and intra- gene pool diversities using Arlequin version 3.5.1 with 1,000,000 markov chain and 100,000 burning steps (Excoffier and Lischer, 2010).

Genome-Wide Association Study

Gummosis disease severity data from 195 peach accessions were used for GWAS based on mixed linear model (MLM) (Segura et al., 2012), fixed and random model circulating probability unification (FarmCPU) (Liu et al., 2016), and Bayesian-information and linkage-disequilibrium iteratively nested keyway (BLINK) (Huang et al., 2019) using GAPIT 3 in R (Wang and Zhang, 2021). The first three principal components were used as covariates for the population structure and familial relatedness calculation, while the kinship matrix was used to eliminate GWAS false positive. The individual relationships were estimated by using the VanRaden method in the GAPIT3 software (VanRaden, 2008). In each step, the variances were estimated by generalized least-square (GLS), and the *P*-values estimated using the *F*-test. All R scripts for converting data format, estimating phenotype BLUP, and plotting pairwise correlation of LD were coded by our research group, and GWAS programs were performed with default parameters in the GAPIT3 software (Wang and Zhang, 2021). The estimated BLUP values were used as phenotype values in GWAS with the cutoff threshold set as 0.01 and the Bonferroni correction (0.01/total number of markers) to filter the significant markers.

Estimation of Linkage Disequilibrium Block in the Gummosis Disease-Associated Region and Candidate Genes Identification and Their Annotation

A 100-kb region flanking the significant SNPs associated with gummosis disease and located within the high LD regions was

investigated based on the peach genome v2.0 to identify the annotated genes. The annotated gene sequences of the peach genome v2.0 assembly were retrieved from GDR¹ to identify the target genes for the corresponding associated regions. Pair-wise LD between markers was calculated as the squared correlation coefficient (r^2) of alleles using the R package LD heatmap (Shin et al., 2006). We used $r^2 > 0.6$ to filter the candidate regions.

RNA-Seq of the Branch Tissue After Pathogen Inoculation

The 1-year-old branches of the susceptible cultivar “Huyou018” were inoculated with *Botryosphaeria dothidea*. The pathogen was isolated from our own germplasm. The inoculation method was based on a previous study by Gao et al. (2016). The tissue measuring 0.5 cm in a diameter was cut from the lesion area and frozen at -80°C for RNA extraction at 0, 48, 60, 72, and 84 h after inoculation. Total RNA extraction and first-strand complementary DNA (cDNA) synthesis were carried out according to Li's method (Li X.W. et al., 2015). The sequencing libraries were generated using the NEBNext[®]Ultra[™] RNA Library Prep Kit for Illumina[®] (NEB, United States) following the manufacturer's recommendations. Reference genome and gene model annotation files were downloaded from the genome website.² The mapped reads of each sample were assembled by StringTie (v1.3.3b) using a reference-based approach (Pertea et al., 2015). A quantification of gene expression level feature Counts v1.5.0-p3 was used to count the read numbers mapped to each gene. The FPKM of each gene was then calculated based on the length of the gene and reads count mapped to that gene. DESeq2 R package (Love et al., 2014) was used for differential gene expression analysis of pair-wise stages using a model with the negative binomial distribution. The *P*-values were adjusted using the Benjamini and Hochberg's approach to controlling the false discovery rate. Genes with an adjusted *P*-value < 0.05 found by DESeq2 were assigned as differentially expressed. For each sampling stage, three biological replicates were combined for further DEG analysis.

RESULTS

Phenotypic Evaluation of Peach Gummosis Disease and Its Heritability

The average gummosis disease value score from the 2-year dataset displayed continuous normal distribution ranging from 0 to 9. The average disease score value for each accession was highly consistent across the period of 2 years ($r^2 = 0.726$). The estimated BLUP value also had a normal distribution (Figure 1). The results demonstrated the existence of a group of accessions highly resistant to gummosis disease. In 92 (47.6%) accessions, the severity of the disease increased over time from 2018 to 2019. In nine accessions, only one to two lesions were found on the entire trunk and branches with a gummosis disease score value of 1, indicating high resistance. Two highly resistant accessions

¹www.rosaceae.org

²http://www.rosaceae.org/species/prunus_persica/genome_v2.0

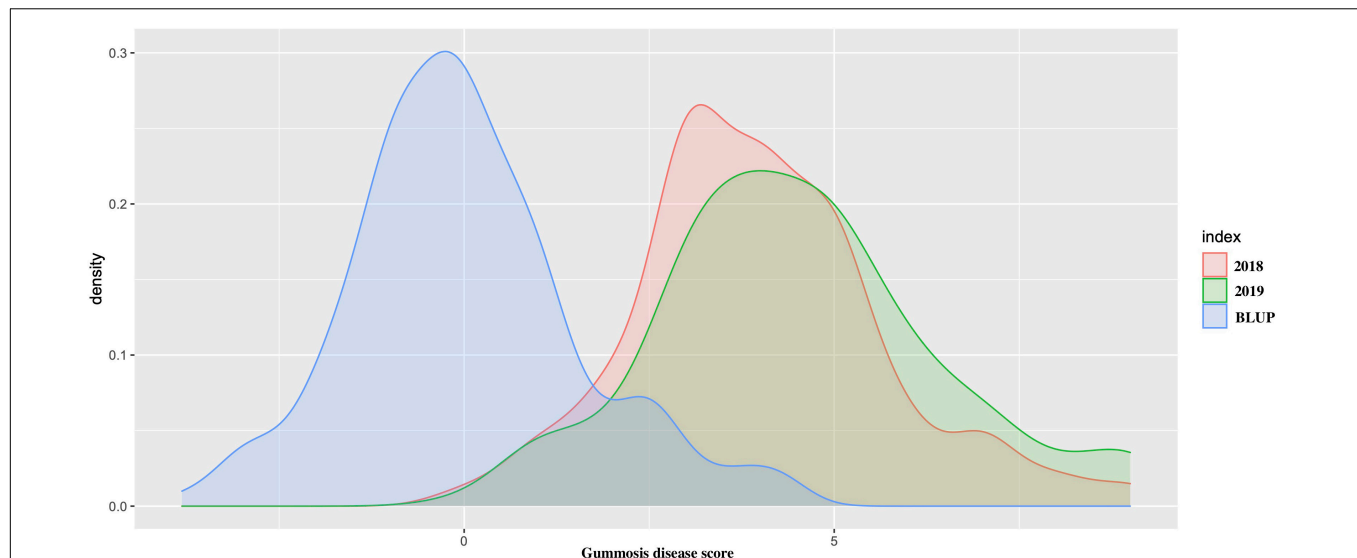


FIGURE 1 | The density distribution of peach gummosis in 195 peach accessions. Disease scale of 0–9 was used, where 0 and 1 represent highly resistant and 9, highly susceptible accessions. The red density distribution is the phenotypic data obtained from 2018; green is the density distribution of the phenotypic data obtained from 2019, and blue represents the density distribution of the best linear unbiased predictions (BLUPs) expressed as estimated breeding values.

“Nan Shan Tian Tao” and “Sunfre” were grown and validated in two field locations. In 2019, 52 accessions had a gummosis disease score under or equal to 3. A total of 134 accessions had a score greater than 3. Of these 31 accessions, including the well-known traditional landraces from several geographic locations, especially north China (“Shenzhou Bai Tao,” “Feicheng Hong Li 6,” “Feicheng Bai Li 10,” and “Taiyuan Shui Mi”) had very severe disease symptoms ranging from 7 to 9.

Among the different peach groups, no significant correlation was observed between gummosis disease and the hairy fruit skin ($r^2 = 0.0004$), fruit flesh color ($r^2 = 0.0004$), fruit shape ($r^2 = 0.002$), geographic origin and domestication history ($r^2 = 0.0006$), and blossoming date ($r^2 = 0.001$). However, the disease severity score in different peach groups separated by geographic origin and domestication history was significantly different. The mean disease severity score of landraces from South China was lower than that of improved accessions from South China and landraces from North China. The mean score in the nectarine group was relatively higher than that in the peach group (Supplementary Figure 2). The broad-sense heritability estimated by multiple years phenotypic data was approximately 70% (Table 1). The narrow-sense heritability estimated by whole genome DNA markers was 73% (Figure 2).

Single Nucleotide Polymorphism Discovery

A total of 1.35 TB of sequence data was generated for the 195 peach genotypes, including 864.45 million reads. The sequencing coverage of at least 1 X was 79.46%. The Q30 ratio, Q20 ratio, and GC content were 85.19, 93.78, and 37.51%, respectively. High-quality reads were aligned with the *Prunus persica* Whole Genome Assembly v2.0 & Annotation

TABLE 1 | Variance components, standard deviations of the variance components, and broad-sense heritability of peach gummosis disease evaluated over 2 years in 195 peach accessions.

| | Accessions | Years | Residuals | Total |
|------------------------|------------|--------|-----------|--------|
| Standard deviation | 1.72589 | 0.5221 | 0.99049 | – |
| Variance | 2.9787 | 0.2726 | 0.9811 | 4.2324 |
| Number of observations | 195 | 3 | – | 1,389 |
| Heritability | 0.70 | | | |

v2.1.³ A total of 9,486,722 SNPs were initially obtained for these genotypes from the SAMTOOLS utility calling (Li et al., 2009). After removing those SNPs with a MAF lower than 0.05 and the missing value higher than 0.07, the remaining set of 145,608 high-quality SNPs was used for further analysis. Among the 145,608 SNPs, 145,456 SNPs (99%) covered all eight chromosomes. The largest number of high-quality SNPs was found on Chromosome 1 (30,358 SNPs), followed by Chromosome 6 (20,173 SNPs); whereas, the smallest number of SNPs was found on Chromosome 8 (13,718 SNPs). The distribution of SNPs on each chromosome was largely consistent with the physical length of the corresponding chromosome.

Population Structure, Genetic Diversity, and Linkage Disequilibrium

The observed heterozygosity per individual ranged from 0.068 to 0.332 with a mean of 0.19 (Supplementary Table 2). The highest value was observed for accession “Hu Zhen 43,” while the lowest value was observed for the traditional *Prunus. Ferganensis*, “Mo Yu 8.” The average value of observed heterozygosity of all SNPs was 0.25. The highest value was observed on Chromosome 4

³https://www.rosaceae.org/species/prunus_persica/genome_v2.0.a1

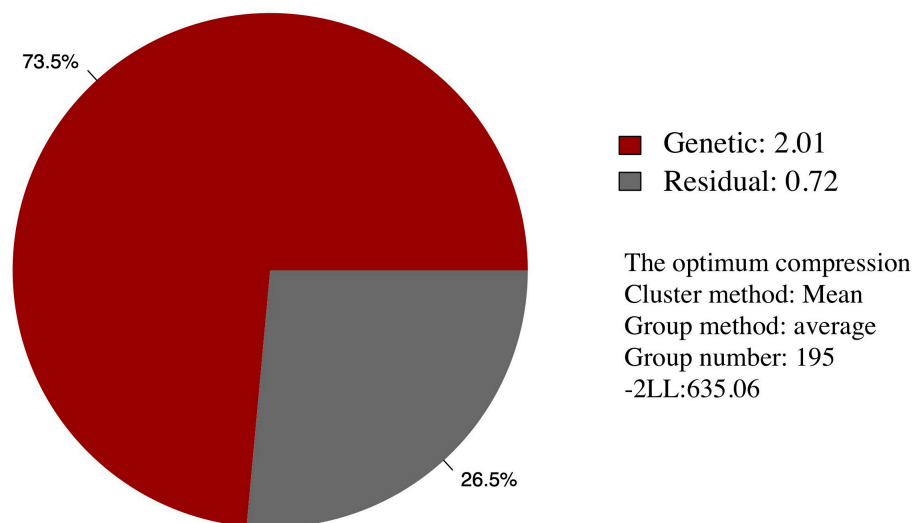


FIGURE 2 | Narrow-sense heritability of the resistance of peach gummosis disease calculated by whole genome SNP markers. “Genetic” and “Residual” means estimated genetic and residual variance in the mixed linear model. The optimum compression information indicates the optimal algorithm to calculate the group kinship matrix, the optimal clustering algorithm, the optimal number of groups in the compress mixed linear model. “-2 LL” is the abbreviation of -2 multiply likelihood value, which means the level of model fitting.

(0.23), and the lowest value was observed on Chromosome 5 (0.15) (**Supplementary Table 4** and **Supplementary Figure 3**).

The geographical origin of the selected accessions could be located at three different continents and 19 provinces in China (**Figure 3A**). A phylogenetic dendrogram using the neighbor-joining method clustered 195 accessions into three major groups, mainly according to fruit shape, hairy vs. glabrous fruit skin, pedigree, geographical origin, and domestication history (**Figure 3B**, **Supplementary Table 1**, and **Supplementary Figure 4**). The first major group was composed of 79 accessions and further divided into three subgroups. The first subgroup G1-1 was marked as the flat peach group, with 10/12 accessions being flat peaches. The accessions of the other two subgroups were closely related to the founder “Shanghai Shui Mi” used in peach breeding programs of China and Japan. One of the most famous cultivars, “Yu Lu,” clustered with the primitive cultivars originating from Shanghai, and most of the Japanese cultivars clustered with “Bai Hua Shui Mi.” The second major group had 30 cultivars and was marked as the traditional landrace group, which included those cultivars carrying special traits, such as red flesh, extremely firm texture, and extremely low chilling requirement. The third major group was composed of 86 accessions and was further divided into two subgroups. The first subgroup G3-1 included 18 peaches and 38 nectarines. In this study, 84% (38/45) of nectarines were clustered in this group. It is noticeable that most of the accessions in this subgroup were characterized as early or very early blossoming. The second subgroup, G3-2, included 27 peaches and three nectarines, and was mainly composed of the accessions with early maturity time. Based on the phylogenetic dendrogram, first approximation of population structure was obtained by using PCA for the complete set of SNPs (**Figure 3C**). The first two principal components explained 43.46% of the total genotypic diversity. The stratification pattern

was highly consistent with NJ hierarchical clustering. All 145,456 SNP markers were employed to estimate the LD extent across the three major groups. The average value of r^2 was 0.269 in G1, 0.133 in G2, and 0.218 in G3. The LD value decreased with distance between the markers in all groups. The level of LD value in G1 was higher than that in G2 and G3. The average value of r^2 dropped below 0.2 at around 30 kb in G2 and 150 kb in G3 (**Figure 3D**).

Based on the population structure, the genetic variation among three major groups was estimated. The result showed that G2, which was a landrace group, had the highest values of F_{is} , π , and haplotype diversity, while the observed heterozygosity value of G2 seemed significantly lower than that of the other groups. The statistical analysis of haplotypes showed that the number of haplotypes and unique haplotypes of G3 was higher than that of G1 and G2 (**Table 2**). The AMOVA revealed that 12.86% of the total variation was found among groups, while the rest of the variation (87.14%) was within groups (**Table 3**). The pairwise genetic differentiation (F_{st}) was highest (0.0909) between G1 and G2 and the lowest (0.0494) between G2 and G3.

Genome-Wide Association Study for Gummosis Disease

Three statistical models were used for GWAS to detect the associated genomic regions with gummosis disease using 145,456 SNP markers and the estimated BLUP values. No significant locus was identified by the MLM model. Five SNPs on five chromosomes were identified as significantly associated with peach gummosis disease (**Figure 4**, **Table 4**, and **Supplementary Figure 5**). Three SNPs were detected by FarmCPU and three by BLINK. Among the five SNPs, rs285957 at about 28 Mb on Chromosome 6 was simultaneously detected both by

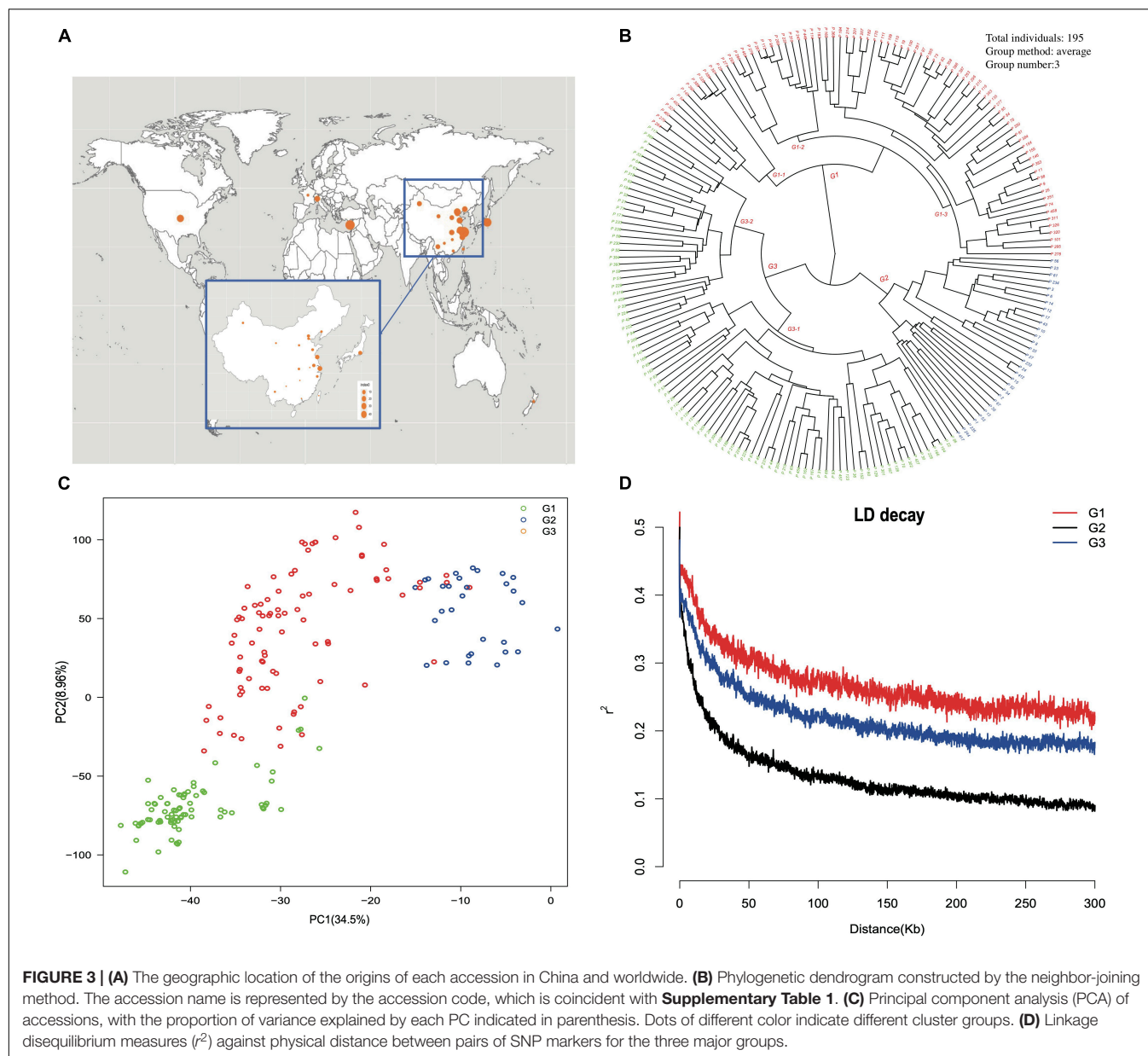


TABLE 2 | The genetic diversity estimated of three major group.

| Group ID | No. of individual | Obs het | Fis | π | No. of haplotype | No. of unique haplotype | Haplotype diversity |
|----------|-------------------|---------|--------|--------|------------------|-------------------------|---------------------|
| G1 | 79 | 0.3653 | 0 | 0.2852 | 13,175,146 | 161,332 | 0.2852 |
| G2 | 30 | 0.2418 | 0.2544 | 0.3256 | 5,003,220 | 161,607 | 0.3255 |
| G3 | 86 | 0.3533 | 0 | 0.3024 | 14,342,564 | 166,600 | 0.3024 |

Obs Het represented observed heterozygosity. Fis indicated inbreeding coefficient. π indicated nucleotide diversity.

FarmCPU and BLINK methods with the allelic effect of 0.64. The phenotypic variation explained by a single SNP varied from 6.28 to 19.85%, and from 6.98 to 17.94% based on FarmCPU and BLINK, respectively. The variance explained by the SNP rs285975 was different under the two models. There were significant phenotypic differences caused by four SNPs in the

different genotypes. The BLUP value of allele “T” at rs22118_C/T, rs142398_C/T, and rs285957_T/G was significantly higher than that of allele “C” and “G,” especially at rs22118. The value of allele “A” at rs191998 was higher than for “T” (Figure 5). It is noticeable that the nine highly resistant accessions carried the same genotype “GG” at rs285957.

TABLE 3 | Analysis of molecular variance of the genetic differentiation among and within three major groups of 195 accessions.

| Source of variation | d.f. | Sum of squares | Variance components | Percentage of variation | P-value |
|---------------------|------|----------------|---------------------|-------------------------|---------|
| Among groups | 2 | 479,575 | 1,885 | 12.86 | 0.001 |
| Within groups | 387 | 4,944,323 | 12,776 | 87.14 | 0.001 |
| Total | 389 | 5,423,898 | 14,661 | | |

Analysis of Differentially Expressed Genes Related to Gummosis Disease at Different Pathogen Inoculation Stages

The fifteen transcriptome sequencing profiles (5 sampling times \times 3 replications) generated a total of 124.91 Gb high-quality data with Phred Quality score or Q30 of 93.43%. The total clean reads for each sample ranged from 43.19 to 70.43 million (**Supplementary Table 4**). The proportion of total mapped reads to the peach reference genome v2.0 accounted for 96.02–97.43%. Of these, the properly mapped reads accounted only 88.31–92.98%. The highest number of DEGs was observed

during the first 48 h after inoculation, including the up- and downregulation of 6,451 and 6,592 genes, respectively. The lowest number of DEGs was observed in 72 vs. 84 h after inoculation, including the up- and downregulation of 2,924 and 2,212 genes, respectively.

The functional annotation of DEGs discovered was performed using gene ontology (GO) functional classification and enrichment analyses. The results showed that three biological process (BP) and six molecular function (MF) terms enriched within 48 h after inoculation, including a response to biotic stimulus (GO:0009607). We also identified a considerable number of DEGs from the functional groups of carbohydrate metabolic process (GO:0005975) in “48 vs. 60 h” and “60 vs. 72 h” comparison (**Supplementary Table 5**). KEGG Pathway enrichment analysis of the DEGs obtained from pairwise comparisons showed that most DEGs were involved in two pathways “Cysteine and methionine metabolism” and “Ribosome” among 0 vs. 48 h. The other three pathways “Flavonoid biosynthesis,” “plant-pathogen interaction,” and “plant hormone signal transduction,” which might be highly correlated with

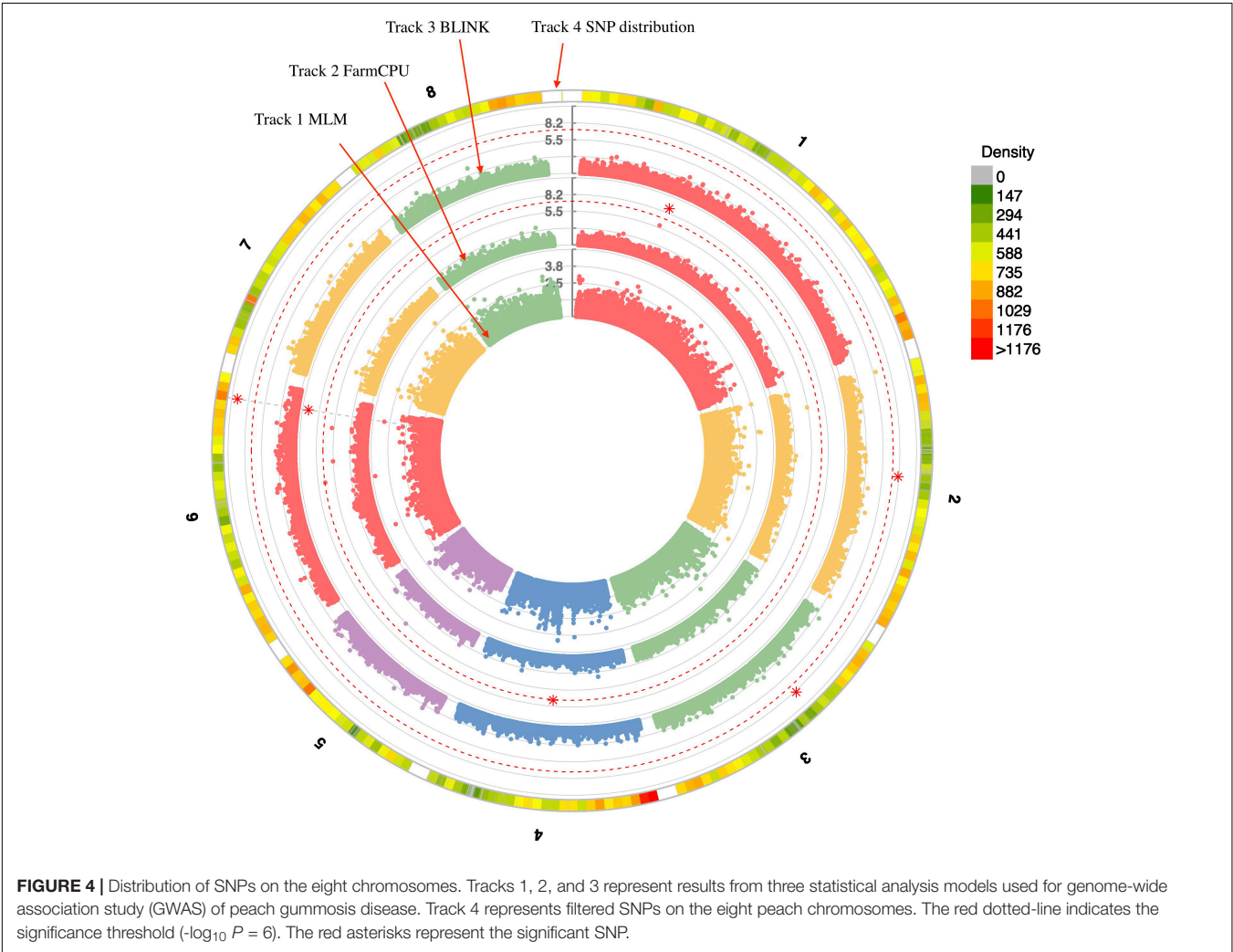
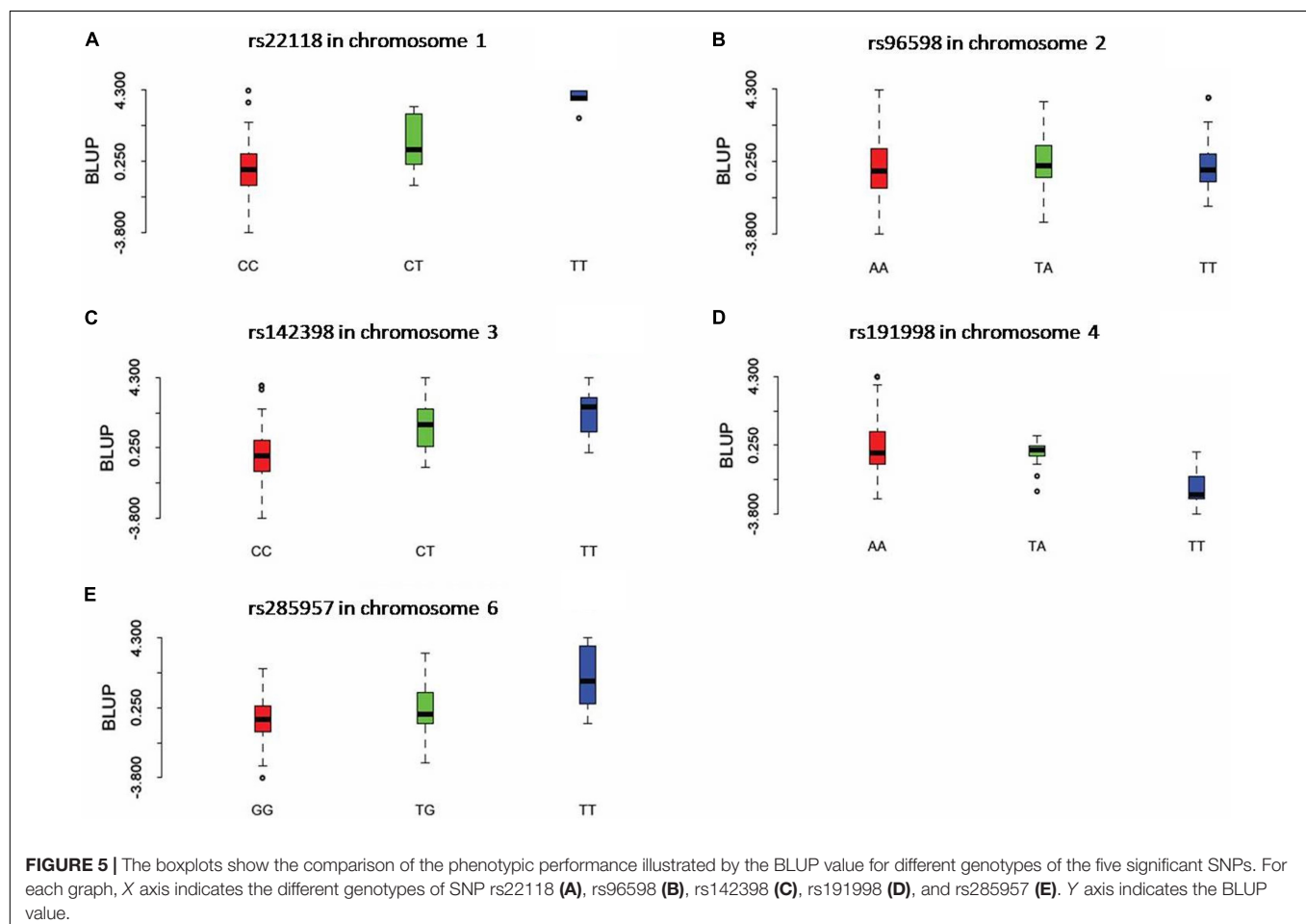


TABLE 4 | A summary of significant SNPs consistently associated with gummosis disease in peach accessions.

| Model | SNP ID | Chromosome | Position | P-value | MAF | FDR | Allelic effect | Variance explained |
|---------|----------|------------|------------|----------|----------|-------------|----------------|--------------------|
| FarmCPU | rs22118 | 1 | 13,801,843 | 9.95E-10 | 0.061538 | 7.24E-05 | 1.113248981 | 19.85 |
| FarmCPU | rs191998 | 4 | 12,480,256 | 5.46E-08 | 0.092308 | 0.002648813 | -0.732419028 | 19.03 |
| FarmCPU | rs285957 | 6 | 28,139,324 | 9.15E-11 | 0.294872 | 1.33E-05 | 0.640842547 | 6.28 |
| BLINK | rs96598 | 2 | 13,388,877 | 7.73E-09 | 0.464103 | 0.000375161 | NA | 6.98 |
| BLINK | rs142398 | 3 | 9,810,975 | 3.20E-09 | 0.089744 | 0.000233144 | NA | 17.94 |
| BLINK | rs285957 | 6 | 28,139,324 | 8.27E-11 | 0.294872 | 1.20E-05 | NA | 17.41 |

FDR in the head row refers to "FDR. Adjusted P-values."

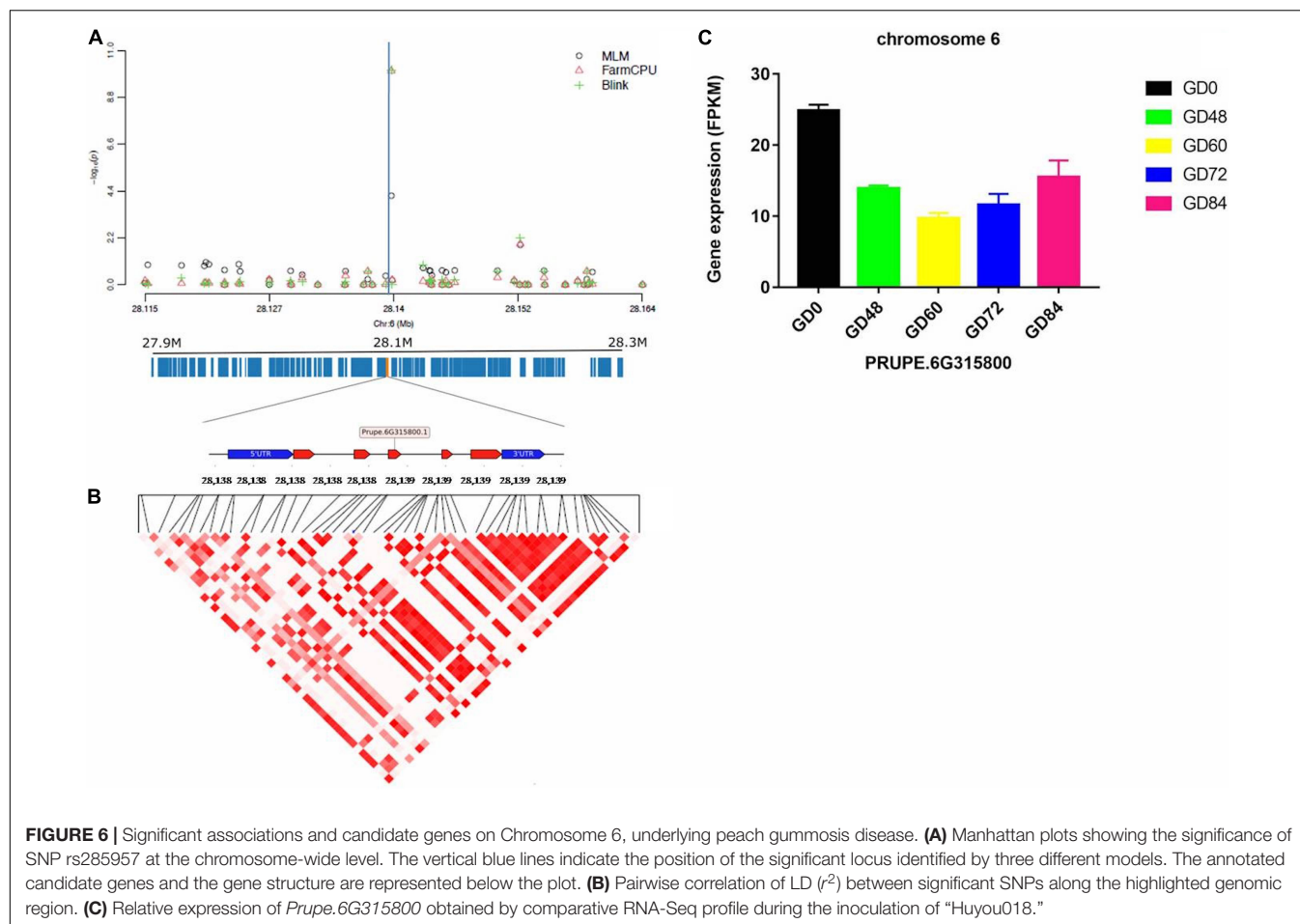


pathogen infection, were also significantly enriched (Supplementary Table 6).

Linkage Disequilibrium Block in the Gummosis Disease-Associated Genomic Regions and Predicted Candidate Genes

The linkage disequilibrium (LD) pattern around each identified significant gummosis disease associated SNPs was evaluated by calculating the squared allele-frequency correlation between each pair of these SNPs. The candidate genes for disease resistance/susceptibility were then searched in the genomic regions flanking the associated SNPs. LD analysis revealed a high

pairwise correlation among SNPs within two candidate genes (*PRUPE.2G084800* and *PRUPE.6G315800*) on Chromosomes 2 and 6, respectively. The putative gene on Chromosome 2 harboring the significant SNP rs96598 is *PRUPE.2G084800* encoding galactose oxidase. The other gene located at the 2 kb upstream region in the same chromosome is a transcriptional activator *PRUPE.2G084700*. It has shown a higher expression level than *PRUPE.2G084800* (Supplementary Figure 6). Upon pathogen infection, these two genes showed significant upregulation from 0 to 72 h and then were downregulated. The significant SNP rs285957 located on Chromosome 6 was found in *PRUPE.6G315800* encoding a Dna J domain, which plays an important role in plant biotic stress in *Arabidopsis*



(Wang et al., 2014). The transcript of *PRUPE.6G315800* decreased after pathogen infection and then increased from 60 to 84 h (Figure 6). The third associated SNP rs142398 located on Chromosome 3 was found within the coding region of a leucine-rich repeat receptor-like protein kinase (LRR-RLK) *PRUPE.3G116000*. Pathogen infection dramatically downregulated the expression level of *PRUPE.3G116000* from 0 to 48 h. The log₂Fold Change of the transcript level of *PRUPE.3G116000* was reduced 3.96 times at 48 h after inoculation compared to 0 h (Supplementary Figure 7). The associated SNP rs191998 on Chromosome 4 was located within the putative gene *PRUPE.4G201700*, which showed an expression pattern similar to *PRUPE.3G116000*. The functional annotation of *PRUPE.4G201700* identified it as a histone H2A.1-like protein (Supplementary Figure 8). Additionally, two UDP-glucosyl transferase genes (*UGTs*, *PRUPE.1G169100*, and *PRUPE.1G169200*) were found to be located at the 20 and 31 kb region downstream of the significant SNP rs22118 on Chromosome 1. However, the pairwise correlation was extremely low within the 25 kb region around the significant SNP (Supplementary Figure 9). The expression of the two genes significantly increased on pathogen inoculation until 72 h, which were similar to *PRUPE.2G084700* and *PRUPE.2G084800*.

DISCUSSION

Gummosis disease is one of the most destructive diseases causing severe loss to peach industry worldwide. At present, no orchard management or fungicide effectively controls the disease in the peach field (Beckman et al., 2011). The absence of resistance cultivars to be used as parents and the poorly understood genetic mechanisms of gummosis disease are the major challenges for breeding the resistant cultivars. In addition, the best solution for breeding a resistant cultivar should be coupled with high fruit quality and not just the root stock, as, even if the rootstock is resistant, the cultivars grafted on the rootstock remain susceptible. In the present study, we investigated large-scale germplasm resources, including improved cultivars and traditional landraces, to identify resistant cultivars for parental selection for future breeding programs. Integrating multiple approaches, including multi-model GWAS and RNA-seq, five novel genetic loci associated with gummosis disease were identified. Four genetic loci with SNPs located on Chromosomes 1, 3, 6, and 4 were found to have a significant impact (P -value < Bonferroni threshold) on the disease severity. The candidate genes harboring the significant SNPs (rs142398, rs285957, and rs191998) were mainly identified and validated by comparative RNA-seq. The functional annotation showed that these genes

were highly related to disease response, which further indicates the reliability of our results.

Phenotypic Variation and the Selection of Resistant Sources

Gummosis is a complex disease as it is reported to be affected by multiple external factors, such as age of the tree, pruning methods, orchard management, temperature humidity, as well as the pathogenic type, plant regulators, and the cultivar itself. By comparing the disease severity of plants in the years of 2018 and 2019, we confirmed the results of Mancero-Castillo that the severity of gumming increases with the age of the tree (Mancero-Castillo et al., 2018). Then, to explain the proportion of phenotypic variance determined by genetic factors, we firstly evaluated the heritability of gummosis disease. The high heritability estimated for peach gummosis disease based on multi-years phenotypic data and genome-wide SNP data is strong evidence that the phenotypic variation of the disease is largely due to genetic effect. This, moreover, indicated the possibility of breeding for resistant cultivar to gummosis disease by introducing resistant parents in the breeding program. Also, because of the high heritability and the significant correlation reported among gum exudates, disease severity, and tissue necrosis (Mancero-Castillo et al., 2018), we mainly evaluated the severity of gummosis disease according to the area of the wound lesion on trunks and branches in the same field, but not according to the pathogen inoculation. Our phenotypic results of 195 germplasm sources showed that most accessions were moderately or highly susceptible to gummosis. Nine *Prunus persica* accessions were highly resistant, of which two accessions showed consistency in resistance across different growing conditions. This result is congruent with the previous reports, where the majority of the peach genotypes were susceptible and no sources of complete resistance were identified even though the available peach genetic resources show the presence of resistant genotypes upon investigation in the field after pathogen infection (Beckman and Reilly, 2005; Beckman et al., 2011). One of the resistant cultivars, “Nan Shan Tian Tao,” is a landrace originating from Shenzhen, southern China, a place with high humidity and temperature (the annual average temperature and humidity ranges up to 25 and 68.5%). This cultivar also exhibited high gummosis resistance in Jiaying, Zhejiang. The resistance of “Sunfre” was validated by growing in two different locations with “3 plus 6” replications in total. Our previous research on 8-year-old trees of “Sunfre,” grown in different experimental fields, also exhibited high resistance to gummosis disease with strong tree vigor and smooth clean trunk (Supplementary Figure 10). The consistent and stable resistance of these two cultivars made them more reliable selections as breeding materials. Although a genetic source for resistance to peach fungal gummosis has been reported from the *P. dulcis* cultivar “Tardy Nonpareil,” which is an almond cultivar (Mancero-Castillo et al., 2018), but it may be simpler and more efficient to choose some famous landraces of *Prunus persica* for intraspecific cross so as to introduce specific alleles, increasing the genetic diversity, and selecting resistant progenies. Accessions with low disease severity will not only be the ideal materials

for breeding superior resistant cultivars but also for identifying disease resistance and related genes in future genetic studies.

Several research groups have reported gummosis disease severity in various peach cultivars. However, there are less comparative studies in different peach groups. Zhao et al. (1996) reported that the severity of the disease was lowest in nectarine groups and highest in the flat peach group. Although we did not identify significant differences on gummosis severity in the groups based on hairy fruit skin, blossom time, and fruit flesh color, a major difference was observed in the groups divided by geographic origin. For example the lowest gummosis disease score was observed in the South China landrace group, and the highest gummosis disease score was observed in the North China landrace group. This result was further confirmed by comparing the disease severity among different subgroups clustered using the neighbor-joining method with genome-wide SNP markers (Supplementary Figure 2). Subgroup G1-2, which contains most of the primitive landraces and offsprings originating from Shanghai, was more resistant than the other subgroups. Shanghai is in Southeast China and has high temperatures and humidity during summer. The monthly average temperature and humidity from June to September range from 23.3 to 28.9°C, and 8 to 85%, respectively. It is known as the origin of flavorful honey peach and elite commercial cultivars worldwide. The result that peach genotypes originating from the Shanghai region showed high resistance to gummosis disease, therefore, indicates toward a selective adaption to climate during acclimation or evolutionary history. The pedigree of modern cultivars is another reason for resistant gene inheritance. For instance, the symptoms in “Early Red 2,” an offspring of “Sunfre” were mild compared to “Huyou” nectarines, derived from “Mayfire,” which had severe disease symptoms (Supplementary Table 7).

Population Structure and Genetic Diversity

Population structure obtained from Neighbor-Joining algorithm and PCA was highly consistent and has clustered the accessions according to the domestication history, pedigree, geographic origin, fruit shape, fruit hairy skin, and blossom time. The traditional landraces were clearly separated from the improved accessions, which, as per the previous reports, were obtained by SSR markers and by the GBS-based SNP array (Li et al., 2013; Micheletti et al., 2015; Cao et al., 2019). Because the accessions originating from Shanghai or developed from “Shanghai Shui Mi” are derived from the same parents, they were grouped together in G1. It is remarkable that the two traditional honey peach (“Yu Lu” and “Bai Hua Shui Mi”), which might have been introduced from Shanghai showed close clustering with the primitive landraces from Shanghai. Henceforth, they are not only the ancestors for the elite cultivars in modern peach breeding programs, but also the most popular cultivars in the fresh-eating market due to the favorable aroma, juicy, melting texture, and high sweetness.

The genetic diversity indices provided useful information on genetic diversity of each population. The high level of genetic diversity within groups and a low level of diversity among groups may be due to gene flow and artificial selection (Eltaher et al.,

2018). The low observed heterozygosity and the highest value of *F_{is}*, π , and haplotype diversity among landraces of G2 can be strong indications that there are no gene flow from landrace in the current breeding program. However, the higher observed heterozygosity and lower genetic diversity among accessions of G1 and G3 may be due to artificial selection of favorable morphological traits and narrow genetic bottleneck because most accessions in these two groups were improved cultivars with desirable traits, such as glabrous fruit skin, strong aroma, sweetness, and low acidity. Some accessions have been frequently used as crossing materials (Li et al., 2019). Considering the above results, the understanding of genetic and phenotypic diversity of G2 will be very helpful for introducing new alleles and enlarging the genetic diversity for creative cultivar selection in the future.

Genome-Wide Association Study Model Selection and Quantitative Trait Locus Identification

Optimal statistical models are needed to accurately evaluate the associations between markers and phenotypes. On comparing the results of three statistical models, no significant SNPs were detected by MLM, which is known as a single locus marker testing model for association study. Many studies have shown that multiple loci markers testing models, such as FarmCPU and BLINK, are more powerful for detection of real association signals and have been integrated into GAPIT3 R packages (Wang and Zhang, 2021). These utilize different testing models to select pseudo QTNs as fixed effect in the final estimated model. FarmCPU uses a set of markers associated with a causal gene as a cofactor instead of kinship to avoid overfitting and eliminates confusion between kinship and testing markers iteratively (Liu et al., 2016). The BLINK eliminates the requirement of FarmCPU, which demands that the quantitative trait nucleotides (QTNs) should be evenly distributed in the genome (Huang et al., 2019). The simulation study has shown that the BLINK model is more powerful than the FarmCPU (Wang and Zhang, 2021). In this study, we compared the results from these two multiple loci models (FarmCPU and BLINK) and selected all significant markers from both models as candidate loci for gummosis disease resistance.

Remarkably, at present, there is only one publication on QTL mapping of peach gummosis disease, which identified a locus Botd8 on chimeric linkage groups 6–8 from “UF Sharp” \times (FG \times TNP1260), with the effect on gumming rates ranging from an average of 0.5 ± 0.2 for resistant to 3.4 ± 0.2 for the susceptible trees (Mancero-Castillo et al., 2018). Here, we identified a total of five quantitative resistance loci affecting gummosis disease by multiple GWAS resolution. All five significant SNPs-harboring genomic loci distributed on five chromosomes (1, 2, 3, 4, and 6) are novel and provided high variance explanation. The large allele effect on phenotypic value is a good indication for detecting the favorable resistance alleles in the current population as well as for future populations. The higher number of QTLs identified by GWAS might be due to higher genetic diversity of our germplasm since most accessions were selected based on the previous study of 658 oriental and occidental cultivars (Li et al., 2013). In addition, comparing with previous linkage mapping study using

bi-parental populations, GWAS gave high mapping resolution to narrow down the chromosomal region of candidate QTLs and predict causal genes (Zhang et al., 2016). However, the SNPs found in our study have not yet been validated in multiple bi-parental populations or natural populations, especially elite parents. This means that there is need for validating the SNPs either using KASP or other convenient and effective tool in training populations to identify favorable alleles that can be selected in future marker-assisted parent selection (MAPS) or marker-assisted seedling selection (MASS) breeding programs.

Identification of Gummosis Disease Resistance Loci and Candidate Genes

Several studies have reported that peach has large LD extent, spanning from around 25–50 kb due to its self-compatibility with limited genetic diversity to be used in peach breeding (Li et al., 2013; Micheletti et al., 2015; Cao et al., 2016). In our study, the LD extent seems to be highly dependent on different groups as it ranged from 30 kb (in G2) to 150 kb in G3. This study is similar to the previous reports (Li et al., 2013; Micheletti et al., 2015; Thurow et al., 2020). However, the LD extent detected for G1 was relatively larger than for G2 and G3. This may be because most accessions in this group have originated from Shanghai or derived from “Shanghai ShuiMi.” With the above view in mind, candidate genes within a conservative window size of approximately 100 kb were searched, and their LD level was analyzed. The SNP rs285927 located in a Dna J domain was identified using both FarmCPU and BLINK models. This is a protein, also known as heat-shock protein 40, which belongs to the family of conserved co-chaperones for HSP70s. Plant J-domain proteins have been shown to have diverse functions in stress responses. For example, silencing a soybean type-III nuclear body-localized DnaJ protein GmHSP40.1 enhanced the susceptibility of soybean plants to soybean mosaic virus (Liu and Whitham, 2013). Similarly, the overexpression of tomato chloroplast-targeted DnaJ protein (LeCDJ2) enhanced the tolerance to drought stress and resistance to *Pseudomonas solanacearum* in transgenic tobacco (Wang et al., 2014). However, virulence effector HopI 1, a chloroplast-targeted class-III J protein from *Pseudomonas syringae*, has been shown to suppress both salicylic acid accumulation and host defense responses in *Arabidopsis* (Jelenska et al., 2007). The comparative transcriptome analysis in this study identified 39 differentially expressed genes that were annotated as Dna J domains. Of these, *PRUPE.6G315800* was co-localized at the same region, where significant SNP rs285957 was detected by GWAS. Moreover, its transcript level also decreased significantly after pathogen inoculation. To further ascertain the function of the DnaJ domain gene family in peach, genome-wide identification and characterization combined with transcript analysis and subcellular localization are necessary. Additionally, another putative gene, *PRUPE.6G315700*, encoding the calmodulin-binding-like protein (CBP), which is related to disease resistance against *Pseudomonas syringae* in *Arabidopsis* and tomato, was found at the 7 kb upstream of the significant SNP locus (rs285957) (Chiasson et al., 2005). In peach, it has been reported that exogenous CaCl_2 treatment can increase the content of Ca^{2+} in shoots, prevent the degradation of cell wall polysaccharides,

maintain the stability and integrity of cell wall, and, finally, reduce the severity of gummosis disease (Li M.J. et al., 2015).

The candidate gene *PRUPE.3G116000* harboring the significant SNP rs142398 on Chromosome 3 is the LRR-RLK gene, which belongs to a large gene family of receptor-like protein kinases and actively participates in regulating growth, development, signal transduction, immunity, and stress responses in plants (Liu et al., 2017; Sun et al., 2017). By performing GBS-based bi-parental QTL mapping and GWAS, clusters of candidate nucleotide-binding site-leucine-rich repeat (NBS-LRR) and receptor-like kinase (RLK) were predicted within the QTL region, which was highly associated with *P. capsici* root rot resistance in pepper (Siddique et al., 2019). In peach, 258 LRR-RLKs genes have been found (Sun et al., 2017). Here, we identified a total of 11 SNPs within *PRUPE.3G116000* using genome sequencing data. However, the correlation between the significant SNP (rs142398) and the other SNPs in the LD block around the gummosis disease-associated genomic region was lower. For this reason, the targeted region was traced in the peach genome v2 in GDR.⁴ As a result, ten SNPs on the IRSC Peach 9K and 18 K SNP array located in the coding region of *PRUPE.3G116000* were found. It is worth noting that the haplotype block constructed with the peach IRSC 9 K SNP array by Stijn (Vanderzande et al., 2019) was not found in this region. This indicates that it may not be a conserved gene but a highly diverse region resulting from recombination, selection, or domestication. Therefore, use of multi bi-parental populations or BSA is required to further analyze the association of *PRUPE.3G116000* with peach gummosis disease.

The UDP-glycosyl transferases are a multigenic and highly divergent superfamily of enzymes that are widely found in all living organisms. In plants, many UGTs play important roles in plant defense to biotic and abiotic stresses by glycosylating acceptor molecules, such as anthocyanidins, flavanols, flavonoids, saponins, sterols, terpenoids, phenylpropanoids, and plant hormones, or by detoxifying and deactivating xenobiotics as a pivotal role in plant-pathogen interactions. In wheat and barley, several UGT genes have been reported to enhance their resistance against *Fusariumhead blight* by glycosylating the deoxynivalenol (DON), produced by *Fusarium* fungus to the less toxic D3G, such as the barley *HvUGT13248* and *HvUGT-10W1* (Xing et al., 2016) and the wheat *TaUGT3* (Xing et al., 2016) and *TaUGT6* (He et al., 2020). In peach, 168 UGT genes have been identified and clustered into 16 groups based on the phylogenetic analysis (Wu et al., 2017). Using the RNA-seq technique, six UGTs (*ppa005290 mg*, *ppa023599 mg*, *ppa012496 mg*, *ppa005161 mg*, *ppa025073 mg*, and *ppa016033 mg*), which are mainly involved in the biosynthesis of anthocyanidins and other flavonoids, has been shown to be upregulated by pathogen infection. The tissue around the wounded area changed from green to red and accumulated anthocyanin during disease infection (Gao et al., 2016). It is worth noting that *PRUPE.1G169100* is identical to the *ppa005161* identified in peach genome v1 by Gao et al. (2016). In our study, two UGTs located around

the significant SNP rs22118 were remarkably upregulated by pathogen infection. They have been previously reported to belong to the same cluster and as homologous with *UGT75D1*, *UGT84A1*, *UGT74F2*, and *UGT74F2*, playing a critical role in *Pseudomonas syringae* resistance and involved in salicylic acid glycosylation in *Arabidopsis* (Boachon et al., 2013; Thompson et al., 2016). UGT glycosylation is a critical step in forming glycosylated linalool, which has been reported to have a defensive function in several plant species such as against rice bacterial blight induced by *Xanthomonas Oryzae* PV. *Oryzae* (Xoo) (Antony et al., 2010), citrus canker induced by *Xanthomonas citri* subsp. *citri* (Xcc) (Takehiko et al., 2017) and antibacterial and antifungal activities to Xcc and *Penicillium italicum* in Ponkan mandarin (Shiduku et al., 2013). In peach, *PpUGT85A2* catalyzes the glycosylation of linalool, and the overexpression of this gene increases the production of linalyl- β -D-glucoside (Wu et al., 2019). Here, we did not evaluate the content of anthocyanin or glycosylated linalool in the shoots of different peach cultivars. Although the regulation of anthocyanin biosynthesis and *terpene synthase* genes by UGTs was not investigated, this research provides a new insight into resistance to peach gummosis disease to understand its defense system.

CONCLUSION

The present study is the first to identify multiple genetic factors involved in peach gummosis using GWAS by using a substantial number of peach germplasm accessions. Two highly resistant accessions were detected in the germplasm, which will be useful plant material for resistant cultivar selection in peach breeding programs. Strong evidence was provided on its high heritability both by genotypic and phenotypic data for peach gummosis disease. This indicates that the phenotypic variation of this complex trait is largely determined by genetic control. By integrating the GWAS and RNA-seq analysis, four candidate genes harboring the significant SNPs on chromosomes 2, 3, 4, and 6 and showing significant differential expression were identified. This study enhances our knowledge of the genetic basis of resistance to peach gummosis disease. The associated markers and resistant plant sources can assist a precise breeding to develop breeders in developing higher resistant cultivars to the disease at a faster rate.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: National Center for Biotechnology Information (NCBI) BioProject database under accession number PRJNA746706.

AUTHOR CONTRIBUTIONS

XL and ZY initiated the project, designed the experiments, and selected the core collection. MS, JD, HZ, and XZ collected the

⁴<https://www.rosaceae.org/>

DNA samples. LW, WF, and KC provided some peach accessions. XL, JZ, ZG, MZ, JJ, KG, and HJ scored the gummosis disease. XL and JW analyzed the phenotypic and genotypic data. XL drafted the manuscript. XL, JW, ZY, and ZG wrote and reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by funds from the National Key Research and Development Program of China (2019YFD1000801), Shanghai Science and Technology

Committee Rising-Star Program (19QB1404600), the Outstanding Team Program of Shanghai Academy of Agricultural Science (Grant No. 2022–004), and the Key Project for New Agricultural Cultivar Breeding in Zhejiang Province, China (2021C02066-4).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.763618/full#supplementary-material>

REFERENCES

- Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., et al. (2010). Rice *xa13* recessive resistance to Bacterial Blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* 22, 3864–3876. doi: 10.2307/41059395
- Aranzana, M. J., Decroocq, V., Dirlwanger, E., Eduardo, I., Gao, Z. S., Gasic, K., et al. (2019). Prunus genetics and applications after de novo genome sequencing: achievements and prospects. *Horti. Res.* 6, 1–25. doi: 10.1038/s41438-019-0140-8
- Arruda, M. P., Brown, P., Brown-Guedira, G., Krill, A. M., Thurber, C., Merrill, K. R., et al. (2016). Genome-wide association mapping of fusarium head blight resistance in wheat using genotyping-by-sequencing. *Plant Genome* 9, 1–14. doi: 10.3835/plantgenome2015.04.0028
- Beckman, T. G., and Reilly, C. C. (2005). Relative susceptibility of ornamental peach cultivars to fungal gummosis (*Botryosphaeria dothidea*). *J. Am. Pomol. Soc.* 60, 149–154.
- Beckman, T. G., Reilly, C. C., Pusey, P. L., and Hotchkiss, M. (2011). Progress in the management of peach fungal gummosis (*Botryosphaeria dothidea*) in the southeastern US peach industry. *Mater. Sci. Appl.* 65, 192–200.
- Boachon, B., Gamir, J., Pastor, V., Dean, J. V., Flors, V., and Mauch-Mani, B. (2013). Role in Defense of the Two Glycosyltransferases UGT74F1 and UGT74F2 Against *Pseudomonas syringae*. *Eur. J. Plant Pathol.* 89, 99–102. doi: 10.1007/s10658-014-0424-7
- Britton, K. O., and Hendrix, F. F. (1982). Three species of *Botryosphaeria* cause peach tree gummosis in Georgia. *Plant Dis.* 66:1120. doi: 10.1094/PD-66-1120
- Cao, K., Li, Y., Deng, C. H., Gardiner, S. E., Zhu, G. R., Fang, W. C., et al. (2019). Comparative population genomics identified genomic regions and candidate genes associated with fruit domestication traits in peach. *Plant Biotechnol. J.* 17, 1954–1970. doi: 10.1111/pbi.13112
- Cao, K., Zhou, Z. K., Wang, Q., Guo, J., Zhao, P., Zhu, G. R., et al. (2016). Genome-wide association study of 12 agronomic traits in peach. *Nat. Commun.* 7, 1–10. doi: 10.1038/ncomms13246
- Chiasson, D., Ekengren, S. K., Martin, G. B., Dobney, S. L., and Snedden, W. A. (2005). Calmodulin-like proteins from *Arabidopsis* and tomato are involved in host defense against *Pseudomonas syringae* pv. tomato. *Plant Mol. Biol.* 58, 887–897. doi: 10.1007/s11103-005-8395-x
- 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
- Eltaher, S., Sallam, A., Belamkar, V., Emara, H. A., Nower, A. A., Salem, K. F. M., et al. (2018). Genetic diversity and population structure of F3:6 nebraska winter wheat genotypes using genotyping-by-sequencing. *Front. Genet.* 9:76. doi: 10.3389/fgenet.2018.00076
- Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x
- Fan, W., Zhao, L., Li, G., Huang, J., and Hsiang, T. (2011). Identification and characterization of *Botryosphaeria* spp. causing gummosis of peach trees in Hubei province, central China. *Plant Dis.* 95, 1378–1384. doi: 10.1094/PDIS-12-10-0893
- Faust, M., and Timon, B. (1995). Origin and dissemination of peach. *Hortic. Rev.* 17, 331–379.
- Gao, L., Wang, Y. T., Li, Z., Zhang, H., Ye, J. L., and Li, G. H. (2016). Gene expression changes during the gummosis development of peach shoots in response to *Lasiodiplodia theobromae* infection using RNA-Seq. *Front. Physiol.* 7:170. doi: 10.3389/fphys.2016.00170
- Guo, D. L., Zhao, H. L., Li, Q., Zhang, G. H., Jiang, J. F., Liu, C. H., et al. (2019). Genome-wide association study of berry-related traits in grape [*Vitis vinifera* L.] based on genotyping-by-sequencing markers. *Horti. Res.* 6:13. doi: 10.1038/s41438-018-0089-z
- He, Y., Wu, L., Liu, X., Jiang, P., and Ma, H. (2020). TaUGT6, a Novel UDP-Glycosyltransferase gene enhances the resistance to FHB and DON accumulation in Wheat. *Front. Plant Sci.* 11:574775. doi: 10.3389/fpls.2020.574775
- Huang, M., Liu, X. L., Zhou, Y., Summers, R. M., and Zhang, Z. W. (2019). BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* 8:giy154. doi: 10.1093/gigascience/giy154
- Jarquín, D., Kocak, K., Posadas, L., Hyma, K., Jedlicka, J., Graef, G., et al. (2014). Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC Genomics* 15:740. doi: 10.1186/1471-2164-15-740
- Jelenska, J., Yao, N., Vinatzer, B. A., Wright, C. M., Brodsky, J. L., and Greenberg, J. T. (2007). A J domain Virulence Effector of *Pseudomonas syringae* Remodels Host Chloroplasts and Suppresses Defenses. *Curr. Biol.* 17, 499–508. doi: 10.1016/j.cub.2007.02.028
- 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, X. W., Jiang, J., Zhang, L. P., Yu, Y., Ye, Z. W., Wang, X. M., et al. (2015). Identification of volatile and softening-related genes using digital gene expression profiles in melting peach. *Tree Genet. Genomes* 11:71. doi: 10.1007/s11295-015-0891-9
- Li, M. J., Liu, M. Y., Peng, F. T., and Fang, L. (2015). Influence factors and gene expression patterns during MeJA-induced gummosis in peach. *J. Plant. Physiol.* 182, 49–61. doi: 10.1016/j.jplph.2015.03.019
- Li, X. W., Meng, X. Q., Jia, H. J., Yu, M. L., Ma, R. J., Wang, L. R., et al. (2013). Peach genetic resources: diversity, population structure and linkage disequilibrium. *BMC Genet.* 14:84. doi: 10.1186/1471-2156-14-84
- Li, Y., Cao, K., Zhu, G. R., Fang, W. C., Chen, C. W., Wang, X. W., et al. (2019). Genomic analyses of an extensive collection of wild and cultivated accessions provide new insights into peach breeding history. *Genome Biol.* 20:36. doi: 10.1186/s13059-019-1648-9
- Liu, H. X., Tan, W. P., Sun, G. W., Zhao, Y. T., and Zhu, X. P. (2015). First report of gummosis disease of apricot (*Prunus armeniaca*) caused by *Botryosphaeria obtusa* in China. *Plant Dis.* 99:6. doi: 10.1094/PDIS-09-14-0981-PDN
- Liu, J. Z., and Whitham, S. A. (2013). Overexpression of a soybean nuclear localized type-III DnaJ domain-containing HSP40 reveals its roles in cell death and disease resistance. *Plant J.* 74, 110–121. doi: 10.1111/tpj.12108
- Liu, P. L., Du, L., Huang, Y., Gao, S. M., and Yu, M. (2017). Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. *BMC Evol. Biol.* 17:47. doi: 10.1186/s12862-017-0891-5

- Liu, X. L., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. W. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12:e1005767. doi: 10.1371/journal.pgen.1005767
- Love, M. L., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Mancero-Castillo, D., Beckman, T. G., Harmon, P. F., and Chaparro, J. X. (2018). A major locus for resistance to Botryosphaeria Dothidea in Prunus. *Tree Genet. Genomes* 14:26. doi: 10.1007/s11295-018-1241-5
- Micheletti, D., Dettori, M. T., Micali, S., Aramini, V., Pacheco, I., Linge, C. D. S., et al. (2015). Whole-genome analysis of diversity and SNP-major gene association in peach germplasm. *PLoS One* 10:e0136803. doi: 10.1371/journal.pone.0136803
- Minamikawa, M. F., Norio, T., Shingo, T., Toshihiro, S., Akio, O., Hiromi, K. K., et al. (2018). Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia* Nakai). *Sci. Rep.* 8:11994. doi: 10.1038/s41598-018-30154-w
- Perte, M., Perte, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotech.* 33, 290–295. doi: 10.1038/nbt.3122
- Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5:3. doi: 10.3835/plantgenome2012.05.0005
- Popović, T., Menković, J., Prokić, A., Zlatković, N., and Obradović, A. (2021). Isolation and characterization of *Pseudomonas syringae* isolates affecting stone fruits and almond in Montenegro. *J. Plant Dis. Prot.* 128, 391–405. doi: 10.1007/s41348-020-00417-8
- Segura, V., Vilhjálmsson, B. J., Platt, A., Korte, A., Seren, Ü, 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
- Shiduku, T., Hosokawa-Shinonaga, Y., Tamaoki, D., Yamada, S., and Akimitsu, K. (2013). Jasmonate induction of the monoterpene linalool confers resistance to rice bacterial blight and its biosynthesis is regulated by JAZ protein in rice. *Plant Cell Environ.* 37, 451–461. doi: 10.1111/pce.12169
- Shin, J. H., Blay, S., Mcnney, B., and Graham, J. (2006). LD heatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J. Stat. Softw.* 16, 1–10. doi: 10.18637/jss.v016.c03
- Siddique, M. I., Lee, H. Y., Ro, N. Y., Han, K., Kang, B. C., Solomon, A. M., et al. (2019). Identifying candidate genes for phytophthora capsici resistance in pepper (*capsicum annuum*) via genotyping-by-sequencing-based QTL mapping and genome-wide association study. *Sci. Rep.* 9:9962. doi: 10.1038/s41598-019-46342-1
- Sun, J., Li, L., Wang, P., Zhang, S., and Wu, J. (2017). Genome-wide characterization, evolution, and expression analysis of the leucine-rich repeat receptor-like protein kinase (LRR-RLK) gene family in Rosaceae genomes. *BMC Genomics* 18:763. doi: 10.1186/s12864-017-4155-y
- Takehiko, S., Tomoko, E., Ana, R., Hiroshi, F., Shingo, G., Takakazu, M., et al. (2017). Ectopic accumulation of linalool confers resistance to Xanthomonas citri subsp. citri in transgenic sweet orange plants. *Tree Physiol.* 5:654. doi: 10.1093/treephys/tpw134
- Thompson, A. M. G., Iancu, C. V., Dean, J., and Choe, J. Y. (2016). Structural basis of distinct salicylic acid glucosylation in Arabidopsis thaliana by two homologous enzymes: implications for plant stress response. *FASEB J.* 30, 1142–1143.
- Thurrow, L. B., Gasic, K., Bassols Raseira, M., Bonow, S., and Castro, C. M. (2020). Genome-wide SNP discovery through genotyping by sequencing, population structure, and linkage disequilibrium in Brazilian peach breeding germplasm. *Tree Genet. Genomes* 16:10. doi: 10.1007/s11295-019-1406-x
- Vanderzande, S., Howard, N. P., Cai, L. C., Linge, C. D. S., Antanaviciute, L., Bink, M. C. A. M., et al. (2019). High-quality, genome-wide SNP genotypic data for pedigreed germplasm of the diploid outbreeding species apple, peach, and sweet cherry through a common workflow. *PLoS One* 14:e0210928. doi: 10.1371/journal.pone.0210928
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980
- Wang, G., Cai, G., Kong, F., DeNg, Y., Ma, N., and Meng, Q. (2014). Overexpression of tomato chloroplast-targeted DnaJ protein enhances tolerance to drought stress and resistance to Pseudomonas solanacearum in transgenic tobacco. *Plant Physiol. Bioch.* 82C, 95–104. doi: 10.1016/j.plaphy.2014.05.011
- Wang, J., and Zhang, Z. (2021). GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *Genom. Proteom. Bioinform.* S1672–0229, 00177–7. doi: 10.1016/j.gpb.2021.08.005 in press.
- Weaver, D. J. (1974). A gummosis disease of peach trees caused by Botryosphaeria dothidea. *Phytopathology*. 64:1429.
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 1358–1370. doi: 10.1111/j.1558-5646.1984.tb05657.x
- Wu, B., Gao, L., Gao, J., Xu, Y., Liu, H., Cao, X., et al. (2017). Genome-wide identification, expression patterns, and functional analysis of UDP glycosyltransferase family in peach (*Prunus persica* L. Batsch). *Front Plant Sci.* 8:389. doi: 10.3389/fpls.2017.00389
- Wu, B. Q., Cao, X. M., Liu, H. R., Zhu, C. Q., Klee, H., Zhang, B., et al. (2019). UDP-glucosyltransferase PpUGT85A2 controls volatile glycosylation in peach. *J. Exp. Bot.* 70, 925–936. doi: 10.1093/jxb/ery419
- Xing, L. P., He, L. Q., Jin, X., Chen, Q. G., and Wang, X. E. (2016). An UDP-glucosyltransferase gene from barley confers disease resistance to Fusarium Head Blight. *Plant Mol. Biol. Rep.* 35, 1–13.
- Yu, Y., Fu, J., Xu, Y., Zhang, J. W., Ren, F., Zhao, H. W., et al. (2018). Genome re-sequencing reveals the evolutionary history of peach fruit edibility. *Nat. Commun.* 9:5404. doi: 10.1038/s41467-018-07744-3
- Zhang, H., Zhang, D., Wang, F., Hsiang, T., and Li, G. (2020). Lasiodiplodia theobromae-induced alteration in ROS metabolism and its relation to gummosis development in Prunus persica. *Plant Physiol. Bioch.* 154, 43–53. doi: 10.1016/j.plaphy.2020.05.018
- Zhang, J. P., Song, Q. J., Cregan, P. B., and Jiang, G. L. (2016). Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine Max*). *Theor. Appl. Genet.* 129, 117–130. doi: 10.1007/s00122-015-2614-x
- Zhang, L., Zhang, Q., Yang, P., Niu, Y., and Niu, W. (2019). First report of gummosis disease of sweet cherry caused by Botryosphaeria dothidea in China. *Plant Dis.* 103:12. doi: 10.1094/PDIS-07-19-1418-PDN
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., and Yang, T. L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35, 1786–1788. doi: 10.1093/bioinformatics/bty875
- Zhao, M. Z., Zhou, J. T., Guo, H., and Yu, M. L. (1996). Evaluation of the resistance to gummosis disease of different peach cultivars in the field. *Deciduous Fruits*. 3, 11–12.

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 Li, Wang, Su, Zhou, Zhang, Du, Zhou, Gan, Jin, Zhang, Cao, Fang, Wang, Jia, Gao and Ye. 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.



Influence of Elevated Temperatures on Resistance Against Phoma Stem Canker in Oilseed Rape

Katherine Noel^{1,2*}, Aiming Qi[†], Lakshmi Harika Gajula¹, Craig Padley², Steffen Rietz³, Yong-Ju Huang[†], Bruce D. L. Fitt[†] and Henrik U. Stotz[†]

¹ Centre for Agriculture, Food and Environmental Management, School of Life and Medical Sciences, University of Hertfordshire, Hatfield, United Kingdom, ² LS Plant Breeding Ltd., Cambridge, United Kingdom, ³ NPZ Innovation GmbH, Holtsee, Germany

OPEN ACCESS

Edited by:

Valerio Hoyos-Villegas,
McGill University, Canada

Reviewed by:

Dilantha Fernando,
University of Manitoba, Canada
Zhongwei Zou,
University of Manitoba, Canada

*Correspondence:

Katherine Noel
knoel@lspb.eu

†ORCID:

Bruce D. L. Fitt
orcid.org/0000-0003-3981-6456
Aiming Qi
orcid.org/0000-0002-0784-9520
Henrik U. Stotz
orcid.org/0000-0002-9800-7862
Yong-Ju Huang
orcid.org/0000-0001-6537-5792

Specialty section:

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

Received: 29 September 2021

Accepted: 03 January 2022

Published: 02 March 2022

Citation:

Noel K, Qi A, Gajula LH, Padley C, Rietz S, Huang YJ, Fitt BDL and Stotz HU (2022) Influence of Elevated Temperatures on Resistance Against Phoma Stem Canker in Oilseed Rape. *Front. Plant Sci.* 13:785804. doi: 10.3389/fpls.2022.785804

Cultivar resistance is an important tool in controlling pathogen-related diseases in agricultural crops. As temperatures increase due to global warming, temperature-resilient disease resistance will play an important role in crop protection. However, the mechanisms behind the temperature-sensitivity of the disease resistance response are poorly understood in crop species and little is known about the effect of elevated temperatures on quantitative disease resistance. Here, we investigated the effect of temperature increase on the quantitative resistance of *Brassica napus* against *Leptosphaeria maculans*. Field experiments and controlled environment inoculation assays were done to determine the influence of temperature on *R* gene-mediated and quantitative resistance against *L. maculans*; of specific interest was the impact of high summer temperatures on the severity of phoma stem canker. Field experiments were run for three consecutive growing seasons at various sites in England and France using twelve winter oilseed rape breeding lines or cultivars with or without *R* genes and/or quantitative resistance. Stem inoculation assays were done under controlled environment conditions with four cultivars/breeding lines, using avirulent and virulent *L. maculans* isolates, to determine if an increase in ambient temperature reduces the efficacy of the resistance. High maximum June temperature was found to be related to phoma stem canker severity. No temperature effect on stem canker severity was found for the cultivar ES Astrid (with only quantitative resistance with no known *R* genes). However, in the controlled environmental conditions, the cultivar ES Astrid had significantly smaller amounts of necrotic tissue at 20°C than at 25°C. This suggests that, under a sustained temperature of 25°C, the efficacy of quantitative resistance is reduced. Findings from this study show that temperature-resilient quantitative resistance is currently available in some oilseed cultivars and that efficacy of quantitative resistance is maintained at increased temperature but not when these elevated temperatures are sustained for a long period.

Keywords: phoma stem canker, quantitative resistance, climate change, oilseed rape, temperature-sensitivity

INTRODUCTION

The plant immune system consists of two branches: a primary basal defense response known as the pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI), and a specific effector-triggered immune (ETI) response. PTI recognizes conserved molecules common to classes of microbes while ETI recognizes and responds to effectors produced by pathogens adapted to overcome PTI. Jones and Dangl (2006) originally proposed a Zigzag model to explain the strength and evolution of PTI and ETI; whereas PTI depends on the perception of PAMPs by pattern recognition receptors (PRRs), ETI involves effector recognition by nucleotide-binding leucine-rich repeat receptors (NLRs). Since this model was proposed, advances have been made in understanding plant immunity, revealing its limitations (Pritchard and Birch, 2014). It has recently become clear that PTI and ETI influence one another to generate a comprehensive immune response (Yuan et al., 2021). Further issues arose from the confusion over the classification of resistance against apoplastic pathogens, such as *Leptosphaeria maculans*, as PTI or ETI (Jones and Dangl, 2006; Thomma et al., 2011; Stotz et al., 2014). The classification of effector-triggered defense (ETD) as another form of resistance, in addition to ETI and PTI, was first proposed by Stotz et al. (2014). ETD replaces ETI when extracellular apoplastic pathogens are encountered. Involving different receptors (ETI is triggered by intracellular NLRs), ETD differs from ETI in several aspects. The ETD response is delayed relative to ETI, which is often associated with fast, hypersensitive host cell death. Furthermore, with ETD, the pathogen is not killed and may resume growth following the onset of host senescence, or if the host resistance response is otherwise compromised (Stotz et al., 2014). All these plant immune and defensive mechanisms are influenced by temperature changes (Cheng et al., 2013).

Increased temperature has been linked to more severe phoma stem canker in winter oilseed rape crops. Previous studies have agreed that in seasons experiencing elevated temperatures and increased rainfall, the efficacy of *R* genes is negatively influenced, and the stem canker severity is greater (Huang et al., 2006, 2018; Evans et al., 2008). Cotyledon assays showed clear differences between *R* genes in their resilience to maintain efficacy under elevated temperatures. Less is known about how these *R* genes respond individually to temperature in crops. Some work has been done in determining which months are most significant in affecting phoma severity; Huang et al. (2018) found phoma leaf spotting and stem canker severity to be linked to October and June average temperatures, respectively. This severity analysis did not explore the impact of maximum monthly temperatures. There is some evidence that maximum temperature may influence canker severity. A multiple linear regression analysis on 40 winter oilseed rape field experimental datasets by Evans et al. (2008) indicated that the mean maximum daily temperature and total rainfall (between 15 July and 26 September) produced the best prediction of the start date of the phoma leaf spotting epidemic, which is used to time the spraying of fungicides in autumn in the United Kingdom for all sites and growing seasons included. Maximum daily temperature and rainfall are important in stage one of the model described

by Evans et al. (2008) relating to the date of leaf spotting in autumn. This study relates to stages two and three of the model (date of canker appearance in spring; severity of canker before harvest); for these stages, only temperature and host resistance are important. June is known to be a critical period in the development of the phoma stem canker; the most severe stage of the disease, the crown canker, occurs from May to July (West et al., 2001).

Conclusions drawn from investigations into the response of quantitative resistance at increased temperatures are somewhat conflicting. Huang et al. (2009) found, by analyzing stem cross-sections, the efficacy of quantitative resistance to be reduced when a cultivar with good quantitative resistance was exposed to an elevated temperature of 25°C compared to 15°C. While more severe cankers were observed on the cultivar with “little” quantitative resistance at 15°C, no significant difference between the two cultivars in stem canker severity at the higher temperature was observed, suggesting that temperature modifies the response of quantitative resistance to *L. maculans*. An experiment by Hubbard and Peng (2018) subjected *L. maculans*-inoculated *Brassica napus* cultivars with quantitative resistance to a temperature regime designed to mimic a heatwave, increasing to 32°C daytime temperature for 7 h before decreasing to 18°C for 7 h overnight. No difference in disease severity was found compared to plants grown at a moderate temperature regime of 22°C daytime/16°C overnight; suggesting that quantitative resistance can maintain efficacy at increased temperatures. It remains poorly understood how temperature affects the operation of quantitative resistance.

Here we aimed to determine how quantitative resistance and different *R* genes impacted the severity of the phoma stem canker of winter oilseed rape cultivars/breeding lines in field conditions, specifically in relation to the June maximum temperature. Second, we examined the effect of elevated temperature on the quantitative resistance response in stems during the colonization of stem tissues of the host *B. napus* by the pathogen *L. maculans* to develop stem canker.

MATERIALS AND METHODS

Winter Oilseed Rape Field Experiments

A selection of *B. napus* breeding lines and cultivars with “good” or “little” quantitative resistance and major resistance (*R*) genes *Rlm4*, *Rlm7*, or *LepR3* were used in the field and controlled environment (CE) experiments. Field disease data and weather data were then analyzed to investigate the relationships between canker severity in different cultivars/breeding lines and maximum monthly temperatures throughout the growing seasons. The field experiments were run for three growing seasons (2016/17, 2017/18, and 2018/19) in England and France. There were two sites in 2016/17; Impington, Cambridgeshire, United Kingdom (lat. 52.253824°, long. 0.125801°) (the previous crop was wheat) and Châteauroux, France (lat. 46.5319°, long. 1.3758°) (the previous crop was wheat). There were two sites in 2017/18; Wisbech, Cambridgeshire, United Kingdom (lat. 52.695707°, long. 0.081937458) (the previous crop was pea) and Châteauroux, France (lat. 46.5319°, long. 1.3758°). However, the

crop failed to establish at Châteauroux due to severe flea beetle damage. There were two sites in the United Kingdom in 2018/19: Callow, Herefordshire (lat. 51.994688°, long. -2.756194°) (the previous crop was wheat) and Wisbech, Cambridgeshire (lat. 52.619527°, long. 0.16128927°) (the previous crop was pea).

A total of 12 winter oilseed rape cultivars/breeding lines were selected for field experiments (Table 1). The rationale for the choice of genotypes was to include current cultivars/breeding lines with “good” or “little” quantitative resistance and *R* genes *Rlm4*, *Rlm7*, or *LepR3*. Current United Kingdom cultivars and breeding lines were included in the study to determine if temperature-resilient characteristics are present in commercially available oilseed rape cultivars. Seven of the cultivars/breeding lines have *R* genes with a “good” quantitative resistance background; DK Exception (*Rlm7*), Breeding line A (*Rlm7*), Adriana (*Rlm4*), Jet Neuf (*Rlm4*), Breeding line C (*Rlm4*), Breeding line E (*LepR3*), and Breeding line F (*LepR3*). Three of the cultivars/breeding lines have *R* genes and “little” quantitative resistance backgrounds; Breeding line B (*Rlm7*), Breeding line D (*Rlm4*), and Breeding line G (*LepR3*). Cultivar ES Astrid contains no known *R* genes but has a quantitative resistance background. Cultivar Incentive, which has no known *R* genes and “little” quantitative resistance, was used as a susceptible control. Breeding line C was not included in the first year as it was selected to replace a cultivar that did not establish in the first year of trials due to poor germination of the seed lot. The field experiments were arranged in randomized block designs with two or three replicates. Seeds were sown between late August and early September, at a density of 45 seeds/m² in France and 55 seeds/m² in the UK. Plots were 6 m² for Châteauroux, France (2016/17, 2017/18), Impington (2016/17), and Wisbech, England (2017/18), and 8.6 m² for Wisbech, England (2018/19) and Callow, England (2018/19).

Frequency of Avirulent Alleles in *L. maculans* Populations

To determine the frequencies in the field experiment areas of virulent and avirulent alleles of *L. maculans* toward the *R* genes *Rlm4*, *Rlm7*, and *LepR3*, cultivar Drakkar with no *R* genes and no quantitative resistance was used for sampling *L. maculans* populations. Leaves of Drakkar with phoma leaf spot lesions were taken from Impington in autumn 2015 and Wisbech in autumn 2016 and 2017. *L. maculans* isolates from single pycnidia were obtained as described by Huang et al. (2018). Eight avirulent alleles of different effector genes in each *L. maculans* isolate were identified by inoculating the isolate onto cotyledons of a differential set of cultivars/lines carrying known *Rlm* genes (Huang et al., 2018).

Phoma Stem Canker Severity Assessment on Different Cultivars and Breeding Lines

The severity of phoma stem canker was assessed in July, prior to harvest, and 15 plants were randomly pulled from each plot. The stems were cut at the base, immediately above the root collar and the area of necrotic tissue caused by phoma stem canker in the cross-section was scored using a scale from 0 to 6 (Pilet et al.,

TABLE 1 | Winter oilseed rape cultivars and breeding lines tested in field experiments in 2016/17, 2017/18, and 2018/19 and a controlled environment (CE) experiment.

| <i>R</i> -gene resistance | ‘Good’ quantitative resistance | ‘Little’ quantitative resistance |
|---------------------------|---|--|
| <i>Rlm7</i> | Group 1 DK Exception ¹ , Breeding line A ¹ | Group 2 Breeding line B ¹ |
| <i>Rlm4</i> | Group 3 Adriana ¹ , Jet Neuf ^{1,2} , Breeding line C ^{1*} | Group 4 Breeding line D ^{1,2} |
| <i>LepR3</i> | Group 5 Breeding line E ¹ , Breeding line F ¹ | Group 6 Breeding line G ¹ |
| None | Group 7 ES Astrid ^{1,2} | Group 8 Incentive ¹ , Breeding line H ² |

There were 12 cultivars/breeding lines in the field experiment and 4 cultivars/breeding lines in the CE experiment (one breeding line is different; hence the 13 cultivars/breeding lines in the table).

Cultivars/breeding lines were categorized into eight groups, depending on their combination of *R*-gene and/or quantitative resistance.

Numbers in superscript refer to experiments in which the cultivar/breeding line was used; winter oilseed rape field experiments (1) and CE temperature-sensitivity assay (2).

Breeding lines A, B, C, D, E, F, G, and H are from NPZ and Jet Neuf is an NPZ cultivar. DK Exception is from DEKALB, Incentive is from DSV, Adriana is from Limagrain and ES Astrid is from Euralis.

*Breeding line C was not included in the first year of field experiments.

1998; Delourme et al., 2008); scale 0 = no affected tissue; scale 1 = 1–5% area affected; scale 2 = 6–25% area affected; scale 3 = 26–50% area affected; scale 4 = 51–75% area affected; scale 5 = 76–100% area affected, plant alive, and 6 = 100% area affected, stem broken or plant dead.

Weather Data at Field Sites

The monthly average maximum temperature and total rainfall data were obtained for the five field site locations to assess their effects on phoma stem canker severity. Weather data were obtained from the NASA Langley Research Centre Atmospheric Science Data Centre Surface Meteorological and Solar Energy (SSE) web portal supported by the NASA LaRC POWER Project¹. The average maximum monthly temperature for six months (from September to July) and maximum June temperature were used for analysis to investigate the effect of increased environmental temperatures.

Effects of Temperature on the Growth Rate of Different *L. maculans* Isolates

The growth of *L. maculans* isolates *in vitro* was assessed at both 20°C and 25°C to ensure any differences in phenotype were not caused by differences in pathogen growth rate. The growth rates of *L. maculans* isolates v23.1.3 (*Av1*-4-5-6-7; avirulent against *Rlm4*) and v23.11.9 (*Av1*-5-6-7; virulent against *Rlm4*) were compared at 20 and 25°C; both isolates were derived from a single cross (Balesdent et al., 2001). Mycelial disk inoculum was placed in the center of 9 cm diameter Petri dishes of V8 agar amended with penicillin (20 mg L⁻¹ filter sterilized) and streptomycin (40 mg L⁻¹ filter sterilized). These were then stored for 24 h in darkness at 20°C before transfer to the CE chamber

¹<https://power.larc.nasa.gov/data-access-viewer/>

for further growth. Six replicate Petri dishes were prepared per treatment. Photographs were taken daily over a 5-day period using a NEX-5R camera (Sony) with a 40.4–49 mm lens. Photos were taken from a fixed height and under controlled lighting to reduce image distortion and give color consistency between treatments. Image J software was used to trace the circumference of the isolated colony in each image using the freehand tool. This method was used to provide more accurate results than measuring colony radius with a ruler as isolates of *L. maculans* often grow with an irregular margin rather than a perfect circular perimeter. The dark orange V8 agar provided a clear contrast to the white mycelia allowing the areas of fungal growth to be clearly identified.

Plant Growth and Stem Inoculation for Controlled Environment Assay

The effect of temperature on the quantitative disease resistance and the role of *R* genes during the second symptomless stage of colonization was investigated by the inoculation of the stem bases of *B. napus* young plants with *L. maculans* isolates. Cultivars/breeding lines possessing four different combinations of resistance genotypes were selected; susceptible background with no *R* genes (Breeding line H); “good” quantitative resistance background with no known *R* genes (ES Astrid); susceptible background with *Rlm4* (Breeding line D), and “good” quantitative resistance background with *Rlm4* (Jet Neuf) (Table 1). Two isolates of *L. maculans* were used for inoculation; one avirulent and the other virulent against *Rlm4*; v23.1.3 (*AvrLm4*), and v23.11.9 (*avrLm4*), respectively. Plants were treated with the virulent *L. maculans* isolate to remove any resistance response caused by major *R* gene interactions. Thus, any differences in resistance response observed in these plants were due to differences in the quantitative resistance background. For each of the four treatments (inoculation with avirulent isolate at 20°C, inoculation with avirulent isolate at 25°C, inoculation with virulent isolate at 20°C, and inoculation with virulent isolate at 25°C), 15 plants of each cultivar/breeding line were subdivided into three sub-blocks of five, which were arranged randomly between four trays each containing 15 plants (three sub-blocks). Plants were grown in a 1:1 ratio of MiracleGro and John Innes No 3 compost, in 6 cm × 6 cm wide and 8 cm deep pots, inside CE cabinets at a constant temperature of 20°C (12-h light/12-h dark). Light intensity at plant height was measured to be 320 $\mu\text{mol}/\text{m}^2/\text{s}$. Plants were divided into two groups 24 h prior to inoculation; half were transferred to 25°C, the rest remaining at 20°C. Plants were inoculated after 6-weeks of growth in the CE cabinets by placing a 1 cm² square piece of sponge cloth soaked in 10⁷ ml⁻¹ conidial suspension over a 1 cm cut in the stem, then wrapping with Parafilm to secure it in place.

Image-Based Canker Severity Assessment and Measurement of Plant Health

Assessment of plant health and canker severity was done at 6-weeks post-inoculation. To assess plant health, the following measurements were taken for each plant; leaf number, plant

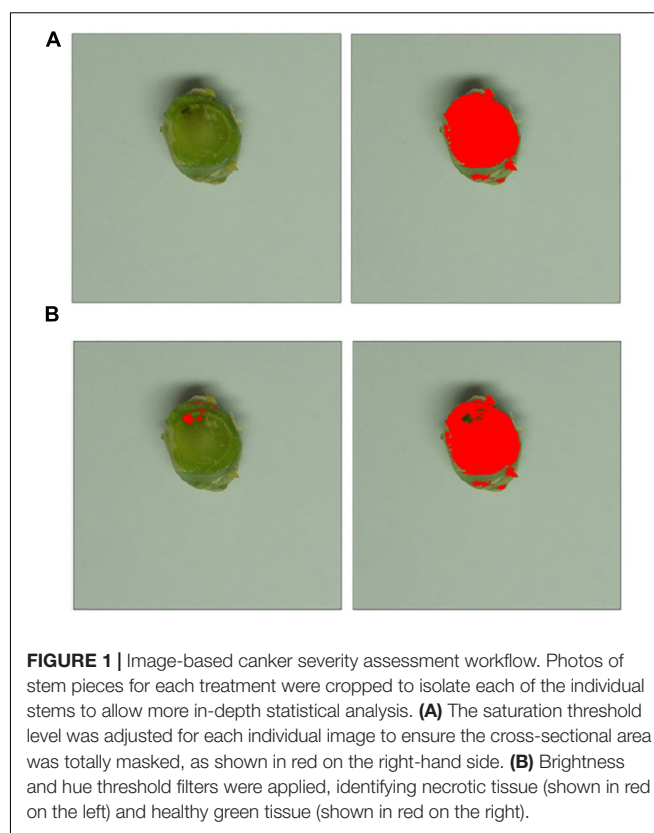


FIGURE 1 | Image-based canker severity assessment workflow. Photos of stem pieces for each treatment were cropped to isolate each of the individual stems to allow more in-depth statistical analysis. **(A)** The saturation threshold level was adjusted for each individual image to ensure the cross-sectional area was totally masked, as shown in red on the right-hand side. **(B)** Brightness and hue threshold filters were applied, identifying necrotic tissue (shown in red on the left) and healthy green tissue (shown in red on the right).

height (stem base to the tip of the longest leaf), and stem thickness (measured with a digital caliper). To assess stem canker severity, 1 cm long pieces of the stem were cut at 1 cm below the inoculation site and photographed as previously described. Photos were then batch-cropped into fifteen photos of stem pieces, each measuring 815 pixels × 815 pixels, to improve accuracy before statistical analysis. These were then analyzed with Image J (Schneider et al., 2012) to determine the percentage area of the necrotic tissue discolored by the disease to assess the severity of the stem canker (Figure 1). Saturation was adjusted for each image to ensure the cross-section of the stem was fully covered in the analysis. Healthy tissue was identified through setting the color threshold parameters, in HSB (hue, saturation, brightness) mode to brightness min 82, hue min 42. Settings for necrotic discolored tissue were brightness max 81 and hue max 41. The Analyze Measure function was then used to measure the pixels in the filtered areas.

Statistical Analysis

The statistical analyses of the data were all done using GenStat statistical software (VSN International, 2020). ANOVA was done to test the effects of cultivars/breeding lines with *R* gene resistance, “good” quantitative resistance, *R* gene resistance with “good” quantitative resistance, or susceptible background on stem canker severity score. ANOVA was also done to test the effects of oilseed rape cultivar/breeding line and temperature on plant height, leaf number, and stem diameter of the oilseed rape cultivars and breeding lines tested. The *post hoc* test with

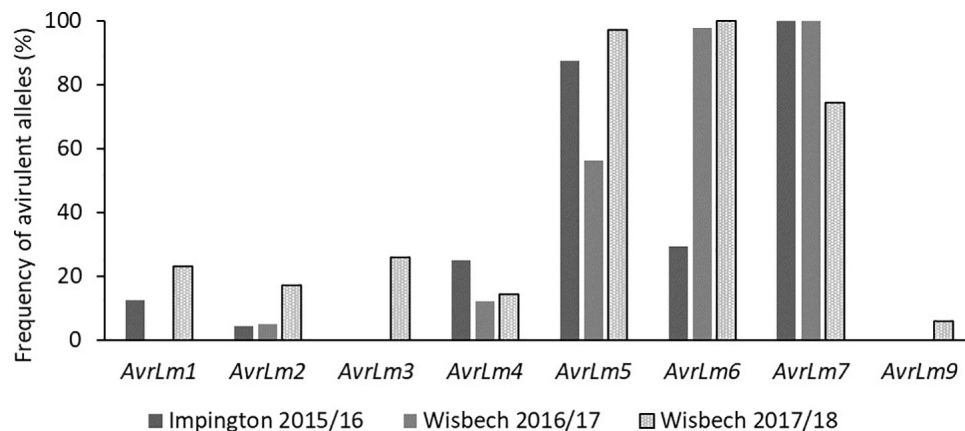


FIGURE 2 | Frequency of avirulent alleles in *Leptosphaeria maculans* isolates taken from field locations in Wisbech and Impington, United Kingdom. To determine the frequencies of virulent/avirulent alleles of *L. maculans* present in the field experiment areas toward the *R* genes *Rlm4*, *Rlm7*, and *LepR3*, cultivar Drakkar with no *R* genes and no quantitative resistance was used for sampling *L. maculans* populations. Leaves of Drakkar with phoma leaf spot lesions were taken from Impington in autumn 2015 and Wisbech in autumn 2016 and 2017 for obtaining *L. maculans* isolates. Avirulent alleles of different effector genes (*Avr*) genes in each *L. maculans* isolate were identified by inoculating the isolate onto cotyledons of a differential set of cultivars/breeding lines carrying known *Rlm* genes (Huang et al., 2018).

Fisher's least significant difference (LSD) calculated at $P = 0.05$ was used to separate the difference between means of treatments. Correlation analysis for canker severity score against mean monthly maximum temperature was done to identify the month with the greatest temperature effect on phoma stem canker severity score. Then, the relationship between the stem canker severity score and the highest maximum temperature recorded in June was analyzed using linear regression. Differences between cultivars/breeding lines were tested using comparative analysis of position and parallelism of linear regression (i.e., linear regression with groups).

RESULTS

Frequency of Avirulent Alleles in *L. maculans* Isolates

The proportions of the avirulent alleles of *AvrLm1*, *AvrLm2*, *AvrLm3*, *AvrLm4*, *AvrLm5*, *AvrLm6*, *AvrLm7*, and *AvrLm9* were assessed in *L. maculans* isolates obtained from leaf lesions taken from experimental locations in Impington (autumn 2015) and Wisbech (autumn 2016 and 2017). The frequency of isolates with *AvrLm7* in Wisbech decreased in 2017/18 (74.3%) compared to 2015/16 and 2016/17 (100%) (Figure 2) and would be expected to decrease further during 2018/19. Thus, cultivars/breeding lines with *Rlm7* would be expected to have good resistance against phoma stem canker in the first year; that would deteriorate during the second and third year of the field experiments as the *avrLm7* races increased in frequency. Most of the isolates tested were found to have the virulent alleles *avrLm1* (87.5% for Impington 2015/16, 100% for Wisbech 2016/17, and 77.1% for Wisbech 2017/18) and *avrLm4* (75% for Impington 2015/16, 87.8% for Wisbech 2016/17, and 85.7% for Wisbech 2017/18) which confer virulence to resistance genes *Rlm1* or *Rlm4*, respectively. Since the effector gene *AvrLm1* is recognized by

the resistance genes *Rlm1* and *LepR3*, cultivars/breeding lines containing *LepR3* would be expected to show severe phoma stem canker at these sites. Similarly, cultivars/breeding lines containing *Rlm4* would also be expected to be susceptible as the effector gene *AvrLm4-7* is recognized by the resistance gene *Rlm4* (Pilet et al., 1998).

Phoma Stem Canker Severity on Different Cultivars and Breeding Lines

The phoma stem canker severity scores for the twelve winter oilseed rape breeding lines/cultivars with different combinations of *R* genes and/or quantitative resistance included in the field experiments in England and France were analyzed. Figure 3 shows the distribution of phoma canker severity scores between cultivars/breeding lines. The cultivar Incentive ("little" quantitative resistance, no known *R* genes) had the greatest canker severity (mean severity score = 3.88), two times greater than that of the cultivar ES Astrid (with quantitative resistance only). Breeding line G ("little" quantitative resistance and *LepR3*) had the smallest average severity score (0.82). The greatest variance in stem canker severity was observed in cultivar DK Exception ("good" quantitative resistance and *Rlm7*) and the smallest in Breeding line D ("little" quantitative resistance and *Rlm4*).

Cultivars/breeding lines with the same *R* gene and "good" or "little" categorization of quantitative resistance were grouped together to allow cross-comparison (Table 2), and the material containing the same *R* genes is color-coded in Figure 3. Fisher's least significant comparison test was done to test for significant differences between *R* genes and quantitative resistance in average stem canker severity. Large differences were seen in *R* gene effects in cultivars/breeding lines with "little" quantitative resistance; cultivars/breeding lines with *Rlm7*, *Rlm4*, *LepR3*, or no known *R* gene were all significantly different from each other. However, in cultivars/breeding lines with quantitative resistance,

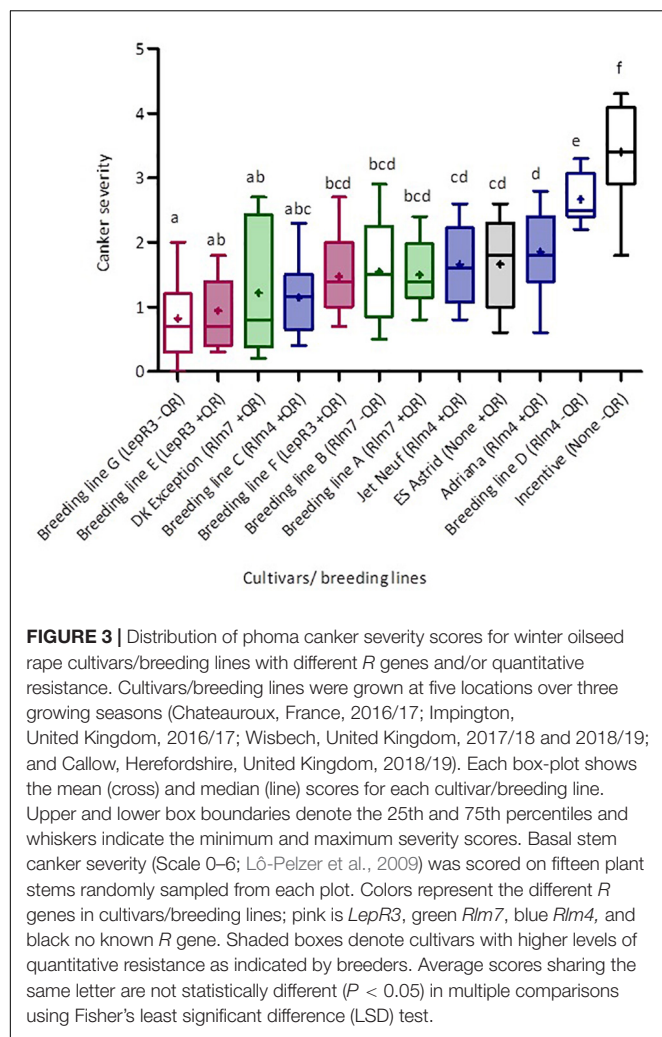


FIGURE 3 | Distribution of phoma canker severity scores for winter oilseed rape cultivars/breeding lines with different *R* genes and/or quantitative resistance. Cultivars/breeding lines were grown at five locations over three growing seasons (Chateauroux, France, 2016/17; Impington, United Kingdom, 2016/17; Wisbech, United Kingdom, 2017/18 and 2018/19; and Callow, Herefordshire, United Kingdom, 2018/19). Each box-plot shows the mean (cross) and median (line) scores for each cultivar/breeding line. Upper and lower box boundaries denote the 25th and 75th percentiles and whiskers indicate the minimum and maximum severity scores. Basal stem canker severity (Scale 0–6; Lô-Pelzer et al., 2009) was scored on fifteen plant stems randomly sampled from each plot. Colors represent the different *R* genes in cultivars/breeding lines; pink is *LepR3*, green *Rlm7*, blue *Rlm4*, and black no known *R* gene. Shaded boxes denote cultivars with higher levels of quantitative resistance as indicated by breeders. Average scores sharing the same letter are not statistically different ($P < 0.05$) in multiple comparisons using Fisher's least significant difference (LSD) test.

there were no *R* gene effects. No significant differences were found for *Rlm7* and *LepR3* cultivars/breeding lines between those with “good” and “little” quantitative resistance. However, for cultivars/breeding lines with *Rlm4* and cultivars/breeding lines with no known *R* gene, significant differences were found between those with “good” and “little” quantitative resistance. In both cases, those with “good” quantitative resistance had a significantly smaller score. Quantitative resistance had the largest protective effect against stem canker caused by *L. maculans* in the absence of *R* genes.

Effect of June Temperature on Canker Severity in Cultivars and Breeding Lines Varying in *R* Gene-Mediated and/or Quantitative Resistance

An initial correlation analysis for the canker severity score and mean monthly maximum temperature was done to identify the month with the greatest temperature effect on phoma stem canker severity score. June was found to have the greatest influence, with a correlation coefficient of $r = 0.33$

TABLE 2 | Fisher's least significance comparison of average canker severity scores for 12 winter oilseed rape cultivars/breeding lines grouped by single *R* gene and quantitative resistance.

| <i>R</i> gene | Quantitative resistance* | | <i>R</i> gene mean |
|-------------------------------------|--------------------------|--------------|--------------------|
| | “Little” | “Good” | |
| <i>Rlm7</i> | 1.53b | 1.36b | 1.421 |
| <i>Rlm4</i> | 2.66c | 1.57b | 1.830 |
| <i>LepR3</i> | 0.82a | 1.20ab | 1.072 |
| None | 3.39d | 1.66b | 2.520 |
| Quantitative resistance mean | 2.074 | 1.433 | |

*Average scores sharing the same letter were not statistically different at $P < 0.05$ in multiple comparisons with Fisher's least significant difference (LSD) test. Values in bold are overall means for genotypes with *R* gene-mediated or quantitative resistance.

(Supplementary Table 1). The regression analysis of the relationship between the greatest recorded June temperature and phoma stem canker severity score in cropping years 2016/17, 2018/19, and 2018/19 showed that there were differences between genotypes (Figure 4). The highest maximum June temperature was 35.97°C in Chateauroux, France (June 22, 2017) and the lowest maximum June temperature was 23.59°C in Wisbech (June 25, 2018).

Groups 1 and 2 (cultivars/breeding lines with *Rlm7* and with “good” or “little” quantitative resistance) both had a positive correlation between the maximum June temperature and phoma stem canker severity. This correlation was stronger in the cultivars/breeding lines with “little” quantitative resistance ($R^2 = 0.85$, $P < 0.025$) than that for the cultivars/breeding lines with “good” quantitative resistance ($R^2 = 0.52$, $P < 0.01$). Groups 3 and 4 (cultivars/breeding lines with *Rlm4*, and “good” or “little” quantitative resistance) showed a much weaker relationship of phoma stem canker severity with maximum June temperature ($R^2 = 0.31$, $P < 0.01$ and $R^2 = 0.26$, $P > 0.05$ respectively). With quantitative resistance, a positive correlation was observed, but with “little” quantitative resistance (Breeding line D), there was no significant correlation. Group 5 (*LepR3* with “good” quantitative resistance) (Breeding lines E and F) and group 6 (*LepR3* and “little” quantitative resistance) (Breeding line G) followed a similar trend to groups 1 and 2; both showed a positive correlation between phoma stem canker score and maximum June temperature, with a stronger correlation in cultivars/breeding lines with “little” quantitative resistance ($R^2 = 0.86$, $P < 0.025$) than those with “good” quantitative resistance ($R^2 = 0.29$, $P < 0.05$). Group 7 (cultivar ES Astrid with “good” quantitative resistance and no known *R* genes) showed no significant correlation between the canker severity score and maximum June temperature ($R^2 = 0.034$). Group 8 (cultivar Incentive with no known *R* genes and “little” quantitative resistance) showed a negative correlation ($R^2 = 0.65$, $P < 0.05$) with reduced phoma stem canker severity at the higher temperatures. Analysis of position and parallelism based on cultivar/breeding line *R* genes compared groups with *R* genes against susceptible cultivar Incentive. Cultivars/breeding lines with *Rlm7* ($P < 0.01$) and *LepR3* ($P < 0.05$) both had significantly

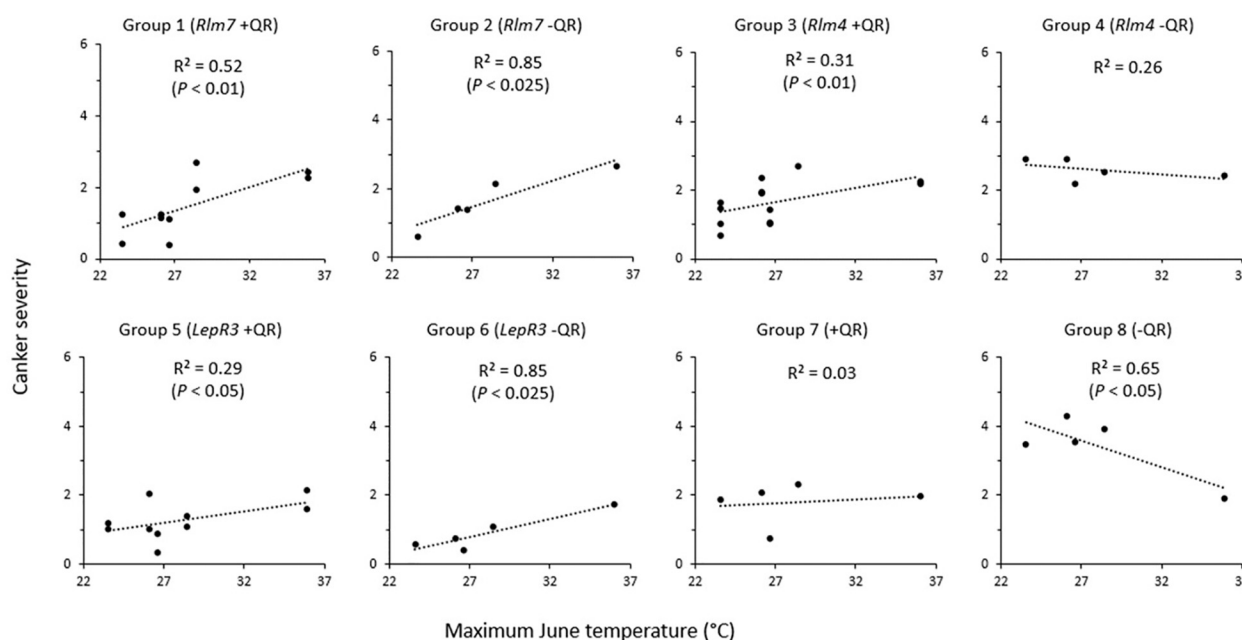


FIGURE 4 | Relationship between phoma stem canker severity of winter oilseed rape cultivars/breeding lines and maximum June temperature. Twelve cultivars/breeding lines (see **Table 1**) were grown in 2–3 replicate blocks over three growing seasons at five locations in Chateauroux, France (2016/17); Impington, Cambridgeshire, United Kingdom (2016/17); Wisbech, Cambridgeshire, United Kingdom (2017/18 and 2018/19); and Callow, Herefordshire, United Kingdom (2018/19). Basal stem canker severity (Scale 0–6; Lö-Pelzer et al., 2009) was scored on 15 plant stems randomly sampled from each plot.

different slopes, but those cultivars/breeding lines with *Rlm4* did not ($P = 0.061$). The intercept was found to be significantly different from Incentive for all cultivars/breeding lines; *Rlm7* ($P < 0.001$), *Rlm4* ($P < 0.05$), and *LepR3* ($P < 0.01$).

Effects of Temperature on the Growth Rate of *L. maculans* Isolates in Culture

The two *L. maculans* isolates used in this study, v23.1.3 and v23.11.9, were grown at 20 and 25°C under CE conditions and measured every 24 h for 5 days to determine the effect of temperature on their growth (**Figure 5**). The growth rates of both isolates were not affected by temperature. The perimeters of the v23.1.3 cultures were slightly larger at 25°C than at 20°C, whereas the perimeters of the v23.11.9 cultures were marginally larger at 20°C. However, neither of these differences were significant.

Effects of Temperature on Phoma Stem Canker Severity in *B. napus* With Quantitative and/or *R* Gene-Mediated Resistance Under Controlled Environment Conditions

The effect of increased temperature, from 20 to 25°C, on canker severity for four winter oilseed rape cultivars/breeding lines with different resistance profiles was determined, 6 weeks following inoculation with *L. maculans* isolates avirulent (v23.1.3) or virulent (v23.11.9) against *Rlm4* (**Figure 6**). *Rlm4* was chosen as the *R* gene in this study as significant differences between cultivars/breeding lines with “little” and “good” quantitative

resistance had previously been observed in the field experiment (**Table 2**). As expected, canker severity was greater when the different plant genotypes were inoculated with the virulent (**Figure 6B**) rather than the avirulent isolate (**Figure 6A**). Cultivar Jet Neuf (“good” quantitative resistance and *Rlm4*) showed the smallest canker severity at both temperatures when inoculated with the avirulent rather than the virulent isolate. No significant difference was seen between the two temperatures for the inoculation with the avirulent *L. maculans* isolate v23.1.3 (11.3 and 24.6% necrosis for 20 and 25°C, respectively); however, the virulent isolate produced significantly greater amounts of necrosis at 25°C (54.1%) compared to 20°C (37.1%). Cultivar ES Astrid (“good” quantitative resistance, no known *R* genes) performed well against isolate v23.1.3 at 20°C with an average necrotic area of 31.8%. However, this cultivar resistance lost efficacy at 25°C, with over twice as much necrotic tissue area (83.7%). When inoculated with the virulent isolate v23.11.9, a significant difference was also seen between the temperatures (63.7 and 84.8% necrotic tissue for 20 and 25°C, respectively). Breeding line D (“little” quantitative resistance and *Rlm4*) also showed a significant difference between the two temperatures when inoculated with the avirulent v23.1.3 isolate (46.3 and 60.3% necrotic tissue area at 20 and 25°C, respectively). When inoculated with the virulent isolate, this temperature effect was reversed, with a statistically significantly greater necrotic tissue area (90.7%) found at 20°C, compared to 69.4% at 25°C. Breeding line H (“little” quantitative resistance, no known *R* genes) did not exhibit any significant temperature effect, although for both isolates the canker severity was slightly less at the

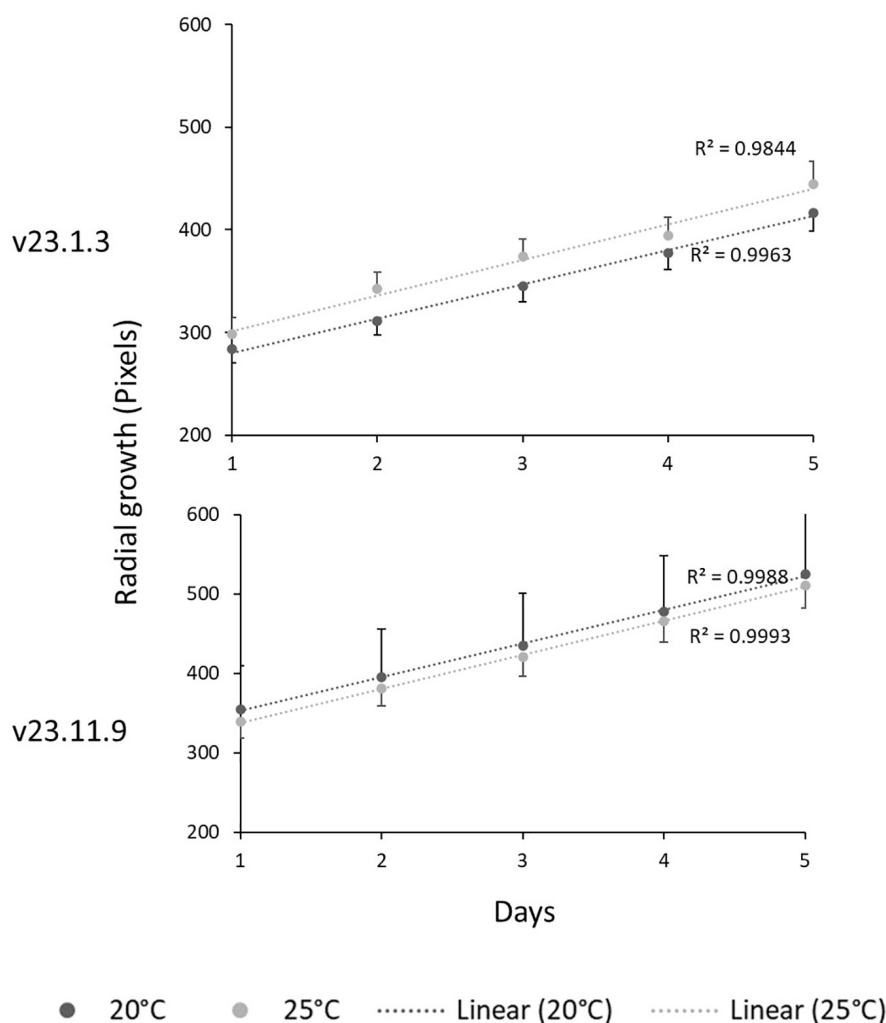


FIGURE 5 | Effects of temperature on the radial growth rates of *L. maculans* isolates v23.1.3 and v23.11.9 at 20 and 25°C. Mycelial disks were transferred from fungal colonies onto V8 media Petri dishes incubated at a constant temperature of 20 or 25°C in darkness. Photographs were taken daily at regular time points for 5 days. The area of fungal growth was analyzed using Image J. An ANOVA showed that temperature had not significantly affected radial growth rate for v23.1.3 ($P = 0.322$) or v23.11.9 ($P = 0.971$). Error bars indicate the standard error of the mean (5 df).

higher temperature. When inoculated with the isolate v23.1.3, the necrotic tissue area was 66 and 58.4% at 20 and 25°C, respectively. When inoculated with the isolate v23.11.9, it was 94.7 and 83.3% at 20 and 25°C, respectively. Both differences were not significant.

Effects of Temperature on Plant Growth Parameters of the Oilseed Rape Cultivars and Breeding Lines Tested

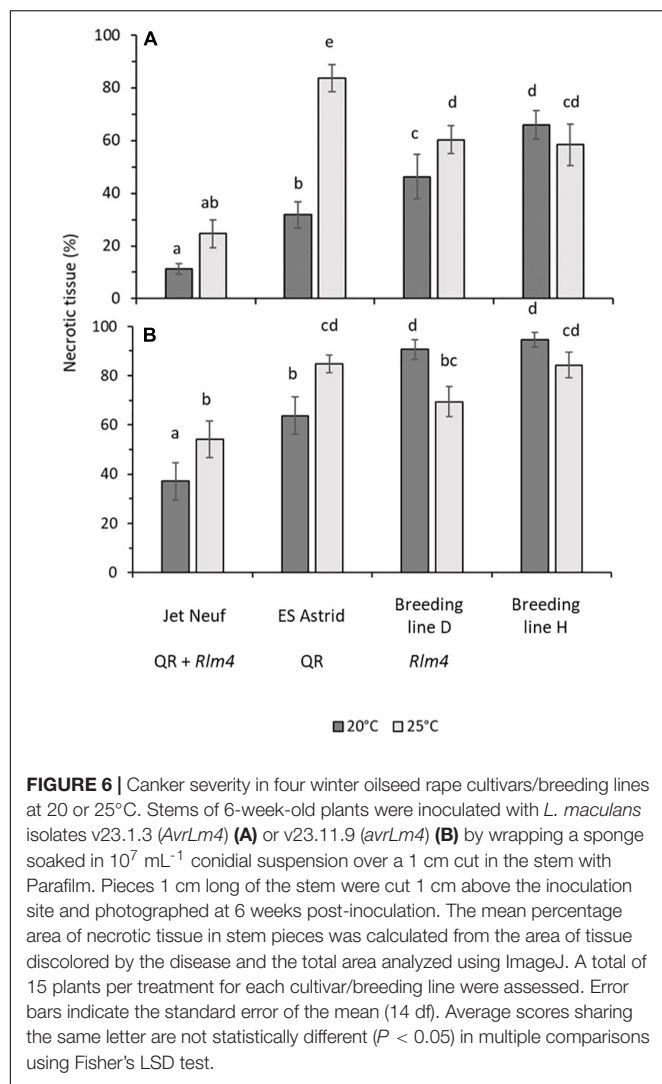
Plant health assessments showed very little difference between the temperatures of 20°C and 25°C (Table 3). Although some differences were seen in stem diameter, these were not significant except for cultivar ES Astrid with a larger stem diameter at 20°C. However, this difference did not affect the image analysis of stem canker because percentages of the area that were necrotic were determined. Cultivar ES Astrid grew better at 20°C than at 25°C;

plants grew taller and had more leaves, but these differences were not significant. Breeding line D grew taller at 25°C than 20°C, significantly when inoculated with isolate v23.11.9; however, significantly more leaves were produced at 20°C when inoculated with this isolate. Breeding line H grew significantly taller at 25°C than at 20°C when inoculated with isolate v23.11.9.

DISCUSSION

Weather Influences Phoma Stem Canker Severity in Genotypes With Different *R* Genes and/or *QR*

This study confirmed the findings of previous work on the effect of temperature on phoma stem canker severity. While the average June temperature showed a correlation with phoma stem canker



severity, a stronger positive relationship was observed for the maximum June temperature recorded and phoma stem canker severity score. This new observation supports the suggestion

that more cultivars will require temperature-resilience to perform successfully in years experiencing high June temperatures, as these are predicted to increase with climate change (Evans et al., 2008; Pullens et al., 2019).

Cultivars/breeding lines with the *R* genes *Rlm7*, *Rlm4*, and *LepR3* responded differently in terms of phoma canker severity to the maximum June temperature. Cultivars/breeding lines with *Rlm7* showed a positive correlation with canker severity increasing with temperature. An alternative explanation could be that *L. maculans* isolates differed between the sites that were tested. Virulent isolates containing *avrLm7* alleles were present in France; the field site Châteauroux experienced the highest maximum June temperature. It is therefore possible that *L. maculans* races, rather than a temperature-sensitive *Rlm7* affected phoma stem canker severity. A recent study of the *Avr* frequencies present in 30 *L. maculans* isolates sampled in Le Rheu, France found 63% had virulence against *Rlm7* (Bousset et al., 2020). Cultivars/breeding lines with *LepR3* had the smallest average canker score, but it also showed a positive relationship with the maximum June temperature.

Breeding line D ("little" quantitative resistance and *Rlm4*) had the second-largest average canker severity; this genotype showed an insignificant correlation between maximum June temperature and canker severity. This could be due to *Rlm4* not being a temperature-sensitive *R* gene; alternatively, a significant proportion of *L. maculans* isolates with virulence against *Rlm4* at the experiment sites, as shown in Figure 2, may have an effect. Analysis of *L. maculans* populations from 13 sites, 11 of which were in the United Kingdom, by Huang et al. (2018) found mean frequencies of *AvrLm4* to be 41%, less than that of *AvrLm7* which was 100%. Within the 30 *L. maculans* isolates sampled in Le Rheu, France, 100% were found to show virulence against *Rlm4* (Bousset et al., 2020). This suggests that *Rlm4* gene-mediated resistance would be rendered at least partially ineffective, explaining the greater canker severity observed in Breeding line D. However, genotypes with *Rlm4* and quantitative resistance had more severe stem canker severity scores at a higher temperature, suggesting that races were "not the end of this story". Collectively, this suggests that

TABLE 3 | Effect of increased temperature from 20 to 25°C on average plant height, leaf number, and total stem diameter of four winter oilseed rape cultivars/breeding lines, inoculated with v23.1.3 or v23.11.9 isolates of *Leptosphaeria maculans*.

| Isolate | | Jet Neuf | | ES Astrid | | Breeding line D | | Breeding line H | |
|----------|--------------------|----------|------|-----------|------|-----------------|------|-----------------|------|
| | | 20°C | 25°C | 20°C | 25°C | 20°C | 25°C | 20°C | 25°C |
| v23.1.3 | Height (cm) | 30.4 | 29.3 | 26.3 | 25.9 | 33.4 | 35.9 | 27.1 | 27.9 |
| | Leaf number | 6.3 | 6.4 | 4.9 | 4.8 | 5.9 | 5.6 | 7.4 | 6.9 |
| | Stem diameter (mm) | 4.2 | 4.4 | 4.3* | 3.6 | 4.6 | 4.9 | 5.4 | 5.3 |
| v23.11.9 | Height (cm) | 27.9 | 28.3 | 26.5 | 25.5 | 30.2* | 34.6 | 23.1* | 29.8 |
| | Leaf number | 5.3 | 5.7 | 5.1 | 4.8 | 6.0* | 5.3 | 7.9 | 7.6 |
| | Stem diameter (mm) | 4.3 | 3.8 | 4.1 | 4.2 | 4.7 | 5.0 | 4.4 | 4.9 |

To compare the differences between variables for v23.1.3, use least significant differences (at $P < 0.05$) for between heights = 2.014; for between leaf numbers = 0.576 and for between stem diameters = 0.467. To compare the differences between variables for v23.11.9, use least significant differences ($P < 0.05$) for between heights = 2.558; for between leaf numbers = 0.632 and for between stem diameters = 0.598. *Significant at $P < 0.05$.

more research is needed on the sampling of isolates from experimental fields together with the characterization of their *Avr* gene profiles.

Lower Temperatures Are More Conducive to Canker Development for Susceptible Cultivars and Breeding Lines

A significant negative correlation of phoma canker severity in field experiments with maximum June temperature was seen for cultivar Incentive, which lacks both known *R* genes and quantitative resistance (Figure 4). Susceptible Breeding line H plants also had a greater amount of necrosis in the stem at 20°C compared to 25°C in the CE experiment, although this difference was not significant when tested separately for each *L. maculans* isolate tested. Together, it may be a result of a greater temperature optimum for PTI. Increased temperatures (23–32°C) have been reported to enhance PAMP signaling in *Arabidopsis thaliana*; on the contrary, ETI has a lower temperature optimum of 10–23°C (Cheng et al., 2013).

R Genes Operate in the Stems of Young Plants Under Controlled Environment Conditions

Results from the field experiments suggested that *R* genes are operating alongside quantitative resistance in June to influence the phoma stem canker severity. Through inoculating stems of young plants, any resistance brought about by *R* genes operating in the leaves was circumvented in stems in the CE experiment. As a control, axenic growth of *L. maculans* was monitored at 20 and 25°C, but no difference in radial growth rate was observed (Figure 5), which is not inconsistent with previous publications (Newbery et al., 2020). While the subtle environmental changes like temperature may influence molecular processes in organisms, it is also rational to assume that organisms can compensate for such changes at least over a certain range. The similar growth rates of *L. maculans* at 20 and 25°C reflect this dynamic range and fit with naturally occurring temperatures in June. Data in Table 3 demonstrated that any observed difference in symptom development under both temperature regimes did not result from different growth rates. Furthermore, it has previously been shown that *L. maculans* can cause disease at both 20 and 25°C (Huang et al., 2006).

The *L. maculans* isolate avirulent to *Rlm4* was found to cause significantly less necrotic tissue in stems of a cultivar with *Rlm4* grown at 20°C than that grown at 25°C. This suggests that *Rlm4* has a protective or suppressive effect against the pathogen growth in the stems of young plants. Previous work showed that *R* genes operate in the leaves of young plants during the autumn to prevent leaf spotting (Rimmer and van den Berg, 1992; Fitt et al., 2006). To confirm this hypothesis, more stem inoculation experiments should be done using near-isogenic lines with or without individual *R* genes. However, little work has been done on the operation of *R* genes in stems. There is a need to test more cultivars with different *R* genes using stem

inoculation, ideally to test near-isogenic lines with or without single *R* genes.

Quantitative Resistance May Protect *R* Gene-Mediated Resistance at High Temperatures

Results of field experiments suggested that quantitative resistance may act to reduce the effect of increasing maximum June temperature on the phoma stem canker severity when combined with *R* genes. This correlation between the maximum June temperature and phoma stem canker severity was weaker for *Rlm7* and *LepR3* cultivars/breeding lines with quantitative resistance compared to those with “little” quantitative resistance. When quantitative resistance was present in a cultivar/breeding line (e.g., ES Astrid) with no known *R* genes, no relationship with maximum June temperature was seen. These findings suggested that quantitative resistance shows temperature-resilience in crops and can buffer a plant resistance response against high temperature, maintaining the efficacy of the plant resistance. However, in the CE stem inoculation experiments, cultivars ES Astrid (with “good” quantitative resistance) and Jet Neuf (*Rlm4* with “good” quantitative resistance) were both found to have a significantly smaller amount of necrotic tissue at 20°C than at 25°C. This suggests that, under a sustained temperature of 25°C, the efficacy of quantitative resistance is reduced. This finding is consistent with previous publications on the temperature sensitivity of quantitative resistance under CE conditions (Huang et al., 2009). One possible explanation for the difference in the performance of cultivar ES Astrid between the field and CE experiments could be the period in which the plant is exposed to elevated temperatures. This finding is supported by CE experiments that mimicked heat waves occurring in Canadian Prairies; gradual increases in temperature from a 7-h night-time period at 18°C to reach a 7-h daytime of 32°C did not change quantitative resistance against *L. maculans*, suggesting that quantitative resistance maintains its efficacy when the increased temperature is not sustained for a long period (Hubbard and Peng, 2018). While quantitative resistance appears to provide a mechanism to reduce the effect of elevated temperature, this may be rendered ineffective if this higher temperature is sustained over a long period.

Although cultivars/breeding lines were classified as having quantitative resistance or not, the resistance mechanisms of different cultivars/breeding lines with quantitative resistance may be completely different. Furthermore, other differences in the genetic backgrounds of these cultivars/breeding lines may also influence their response to the environment and impact the severity of phoma stem canker. Thus, there are clear limitations to this study. Nevertheless, in the absence of a set of oilseed rape lines differing only in their quantitative resistance trait loci, this set of genotypes provides a good choice for investigating the effect of temperature on the quantitative resistance response.

It is not known if increased levels of quantitative resistance are linked with a reduction in fitness. A review by Brown (2002) on yield penalties of disease resistance in crops suggested

that plants with good quantitative resistance could suffer from a fitness penalty. The evidence behind this proposal came from the observations of Vanderplank (1984) that quantitative resistance can be lost due to masking by single *R* genes or if not exposed to the pathogen. Quantitative resistance genes could be linked to genes involved in yield, resulting in linkage drag if these resistance genes were to be introgressed. More research is needed in this area to fully understand any potential trade-offs in important traits, such as yield, that may be linked to greater levels of quantitative resistance in oilseed against phoma stem canker.

CONCLUSION

Results from field experiments suggest that temperature-resilient quantitative resistance is currently available in some oilseed cultivars. For example, ES Astrid (“good” quantitative resistance, no known *R* genes) showed no significant correlation between canker severity score and maximum June temperature. However, ES Astrid had significantly smaller amounts of necrotic tissue at 20°C than at 25°C when inoculated with both virulent and avirulent *L. maculans* isolates under CE conditions. We suggest that the efficacy of quantitative resistance is maintained at increased temperature but not when these elevated temperatures are sustained for long periods of time under CE conditions.

The effectiveness of *Rlm4* mediated resistance in the stem also appears to be reduced when plants are subjected to a prolonged elevated temperature of 25°C. Significantly more necrotic tissue was found at 25°C than 20°C after Breeding line D (“little” quantitative resistance and *Rlm4*) was inoculated with an avirulent *L. maculans* isolate. The reverse was seen when the same line was inoculated with a virulent *L. maculans* isolate. However, in Jet Neuf (“good” quantitative resistance and *Rlm4*) there was no significant difference in the amount of necrotic tissue between the two temperatures, when inoculated with an avirulent isolate. Therefore, in years experiencing warmer summers, as have been predicted to result from climate change in the United Kingdom, a combination of temperature-resilient *R* genes and a good quantitative resistance background will be required to protect oilseed crops from phoma stem canker.

Furthermore, the results of the CE experiments show that both quantitative resistance and *R* gene resistance operate in the stem by either preventing or suppressing the growth of *L. maculans*, subsequently reducing stem canker severity. This is important to growers as yields can be significantly reduced by phoma stem canker developing in the summer months. There may be scope to reduce this damage in the future by assessing new cultivars to determine their level of stem resistance and ability to maintain resistance at elevated temperature, using stem base inoculation methods in CE assays. However, it would be advised that the temperatures in these assays would be set to simulate the types of heatwaves forecast to become more common as global warming advances. Temperatures should be set to fall at night rather than maintain a constant temperature throughout.

From this study, it could be suggested that *Rlm4* is a weaker, yet more temperature-resilient, *R* gene compared to

Rlm7 and *LepR3*. For cultivars/breeding lines with *Rlm4*, significant differences were found between those with “good” and “little” quantitative resistance. However, it must be remembered that high frequencies of virulent isolates with *avrLm4* alleles were found in two of the field experiment locations. Cultivars/breeding lines with *Rlm4*, with “good” or “little” quantitative resistance, showed a much weaker relationship of phoma stem canker severity with maximum June temperature compared to cultivars and breeding lines with *Rlm7* and *LepR3*.

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.

AUTHOR CONTRIBUTIONS

YH, HUS, SR, CP, LG, and KN: experimental design. AQ and KN: statistical analysis. KN and LG: lab work. KN and CP: field work. BDLF and HUS: project administration. HUS: supervision. KN and HUS: writing – original draft. KN, HUS, YH, AQ, and BDLF: writing – review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

We gratefully acknowledge those funding this project; the UK Biotechnology and Biological Sciences Research Council and the Knowledge Transfer Network (BB/N503848/1) (Project reference: 1662042), the UK Biotechnology and Biological Sciences Research Council (BBSRC, M028348/1, and P00489X/1), the Innovate UK (102100 and 102641), AHDB Cereals & Oilseeds (RD-2140021105), and the Chadacre Agricultural Trust and the Perry Foundation. Senior author and principal supervisor of this project, HUS was also supported by BBSRC projects BB/R019819/1 and BB/V01725X/1.

ACKNOWLEDGMENTS

We thank Marie-Hélène Balesdent at the French National Institute for Agriculture, Food, and Environment (INRAE) for providing the *L. maculans* isolates v23.1.3 and v23.11.9.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.785804/full#supplementary-material>

REFERENCES

- Balesdent, M. H., Attard, A., Ansan-Melayah, D., Delourme, R., Renard, M., and Rouxel, T. (2001). Genetic control and host range of avirulence toward *Brassica napus* cultivars Quinta and Jet Neuf in *Leptosphaeria maculans*. *Phytopathology* 91, 70–76. doi: 10.1094/PHYTO.2001.91.1.70
- Bousset, L., Ermel, M., and Delourme, R. (2020). A *Leptosphaeria maculans* set of isolates characterised on all available differentials and used as control to identify virulence frequencies in a current French population. *BioRxiv [Preprint]* doi: 10.1101/2020.01.09.900167
- Brown, J. K. (2002). Yield penalties of disease resistance in crops. *Curr. Opin. Plant Biol.* 5, 339–344. doi: 10.1016/s1369-5266(02)00270-4
- Cheng, C., Gao, X., Feng, B., Sheen, J., Shan, L., and He, P. (2013). Plant immune response to pathogens differs with changing temperatures. *Nat. Commun.* 4:2530. doi: 10.1038/ncomms3530
- Delourme, R., Piel, N., Horvais, R., Pouilly, N., Domin, C., Vallee, P., et al. (2008). Molecular and phenotypic characterization of near isogenic lines at QTL for quantitative resistance to *Leptosphaeria maculans* in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* 117, 1055–1067. doi: 10.1007/s00122-008-0844-x
- Evans, N., Baierl, A., Semenov, M., Gladders, P., and Fitt, B. (2008). Range and severity of a plant disease increased by global warming. *J. R. Soc. Interface* 5, 525–531. doi: 10.1098/rsif.2007.1136
- Fitt, B. D. L., Brun, H., Barbetti, M., and Rimmer, S. (2006). World-wide importance of phoma stem canker (*Leptosphaeria maculans* and *L. biglobosa*) on oilseed rape (*Brassica napus*). *Eur. J. Plant Pathol.* 114, 3–15. doi: 10.1007/s10658-005-2233-5
- Huang, Y. J., Evans, N., Li, Z. Q., Eckert, M., Chèvre, A. M., Renard, M., et al. (2006). Temperature and leaf wetness duration affect phenotypic expression of *Rlm6*-mediated resistance to *Leptosphaeria maculans* in *Brassica napus*. *New Phytol.* 170, 129–141. doi: 10.1111/j.1469-8137.2005.01651.x
- Huang, Y. J., Mitrousis, G., Sidique, S., Qi, A., and Fitt, B. D. L. (2018). Combining *R* gene and quantitative resistance increases effectiveness of cultivar resistance against *Leptosphaeria maculans* in *Brassica napus* in different environments. *PLoS One* 13:e0197752. doi: 10.1371/journal.pone.0197752
- Huang, Y. J., Pirie, E., Evans, N., Delourme, R., King, G., and Fitt, B. (2009). Quantitative resistance to symptomless growth of *Leptosphaeria maculans* (phoma stem canker) in *Brassica napus* (oilseed rape). *Plant Pathol.* 58, 314–323. doi: 10.1111/j.1365-3059.2008.01957.x
- Hubbard, M., and Peng, G. (2018). Quantitative resistance against an isolate of *Leptosphaeria maculans* (blackleg) in selected Canadian canola cultivars remains effective under increased temperatures. *Plant Pathol.* 67, 1329–1338. doi: 10.1111/ppa.12832
- Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286
- Lô-Pelzer, E., Aubertot, J. N., David, O., Jeuffroy, M. H., and Bousset, L. (2009). Relationship between severity of blackleg (*Leptosphaeria maculans*/L. *biglobosa* species complex) and subsequent primary inoculum production on oilseed rape stubble. *Plant Pathol.* 58, 61–70. doi: 10.1111/j.1365-3059.2008.01931.x
- Newbery, F., Ritchie, F., Gladders, P., and Fitt, B. D. L. (2020). Inter-individual genetic variation in the temperature response of *Leptosphaeria* species pathogenic on oilseed rape. *Plant Pathol.* 69, 1469–1481. doi: 10.1111/ppa.13236
- Parlange, F., Daverdin, G., Fudal, I., Kuhn, M. L., Balesdent, M. H., and Blaise, F. et al. (2009). *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7* resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change. *Mol. Microbiol.* 71, 851–863. doi: 10.1111/j.1365-2958.2008.06547.x
- Pilet, M. L., Delourme, R., Foisset, N., and Renard, M. (1998). Identification of QTL involved in field resistance to light leaf spot (*Pyrenopeziza brassicae*) and blackleg resistance (*Leptosphaeria maculans*) in winter rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 97, 398–406.
- Pritchard, L., and Birch, P. (2014). The zigzag model of plant-microbe interactions: is it time to move on? *Mol. Plant Pathol.* 15, 865–870. doi: 10.1111/mpp.12210
- Pullens, J., Sharif, B., Trnka, M., Balek, J., Semenov, M., and Olesen, J. (2019). Risk factors for European winter oilseed rape production under climate change. *Agric. For. Meteorol.* 27, 30–39. doi: 10.1016/j.agrformet.2019.03.023
- Rimmer, S., and van den Berg, C. (1992). Resistance of oilseed *Brassica* spp. To blackleg caused by *Leptosphaeria maculans*. *Can. J. Plant Pathol.* 14, 56–66. doi: 10.1080/07060669209500906
- Schneider, C., Rasband, W., and Eliceiri, K. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089
- Stotz, H. U., Mitrousis, G. K., de Wit, P. G. M., and Fitt, B. D. L. (2014). Effector-triggered defence against apoplastic fungal pathogens. *Trends Plant Sci.* 19, 491–500. doi: 10.1016/j.tplants.2014.04.009
- Thomma, B., Nürnberger, T., and Joosten, M. (2011). Of PAMPs and effectors: the blurred PTI-ETI dichotomy. *Plant Cell* 23, 4–15. doi: 10.1105/tpc.110.082602
- Vanderplank, J. E. (1984). *Disease Resistance in Plants*. London: Academic Press, 1984.
- VSN International. (2020). *GenStat for Windows 21st Edition*. Hemel Hempstead: VSN International.
- West, J. S., Kharbanda, P. D., Barbetti, M. J., and Fitt, B. D. L. (2001). Epidemiology and management of *Leptosphaeria maculans* (phoma stem canker) on oilseed rape in Australia, Canada and Europe. *Plant Pathol.* 50, 10–27. doi: 10.1046/j.1365-3059.2001.00546.x
- Yuan, M., Jiang, Z., Bi, G., Nomura, K., Liu, M., Wang, Y., et al. (2021). Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature* 592, 105–109. doi: 10.1038/s41586-021-03316-6

Conflict of Interest: KN and CP are employed by LS Plant Breeding Ltd. SR is employed by NPZ Innovation GmbH.

The authors declare that this study received funding from LS Plant Breeding. The funder had the following involvement in the study: Contribution of breeding lines and field experiments. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

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.

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 Noel, Qi, Gajula, Padley, Rietz, Huang, Fitt and Stotz. 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.

Frontiers in Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

